# Angry or Climbing Stairs? Towards Physiological Emotion Recognition in the Wild

Judith S. Heinisch, Christoph Anderson, and Klaus David
Chair for Communication Technology (ComTec)
University of Kassel
Kassel, Germany
Email: comtec@uni-kassel.de

***Abstract*—Physiological responses to emotions play a vital role in the field of emotion recognition. Machine-learning models implemented in wristbands or wearables, already exploit unique patterns in physiological responses to provide information about humans emotional states. However, such responses are commonly interfered and overlapped by physical activities, posing a challenge for emotion recognition "in-the-wild". In this paper, we address this challenge by investigating new features based on the linear regression line and machine-learning models for emotion recognition. We triggered emotions through audio samples and recorded physiological responses from 18 participants before and while performing physical activities. We trained models with the least strenuous physical activity (sitting) and tested with the remaining, more strenuous ones. For three different emotion categories, we achieved classification accuracies up to 67%. Considering individual activities and participants, we achieve up to 73% classification accuracy, indicating the viability of emotion recognition models and features non-sensitive to interferences caused by physical activities.**

## I. Introduction

Sensing and recognizing emotional states of individuals are one of the main challenges in the field of Human-Computer Interaction. Research in this field is fueled by the idea to enhance computer systems to a state where they can sense, adapt, or even react to the emotional states of their users. For example, advanced driver assistant systems might sense emotional states of drivers to detect risky driving behaviors [1]. Work-related environments may include emotion recognition to support software developers in their productivity and mitigate effects caused by interruptions [2]. Physical and physiological responses to emotions have been investigated to facilitate various applications of emotion recognition [3]. Among others, microphones and cameras have been used to extract speech, facial expressions, or postures for physically-based emotion recognition [4]. Physical responses, however, are subject to suppression and dissimulation as individuals can control facial expressions or the tone of their speech, therefore, confounding emotion recognition systems [5], [6]. Physiological responses to emotions, however, are difficult to control and are affected by physical movement and activity [4].

Approaches that cope with physically-based interferences, for example, provide models designated and trimmed for individual activities [7], or select appropriate machine-learning models for similar interferences [8]. Although these models are quite practical, they are affected by the kind of interference, or the computation complexity is increased as multiple models are required. Therefore, we aim to address the challenge of recognizing emotions *throughout* and *non-sensitive* to physical activities. To investigate physically-based interferences, we carried out an experiment with 18 participants where emotions were elicited while performing physical activities. To force non-sensitivity, we first filtered the recorded physiological signals and then trained three different machine-learning algorithms with the data of the least strenuous activity. The data of other strenuous activities were then used to evaluate and to assess the performances of machine-learning models. We found that the data of features based on the linear regression line of physiological signals facilitate machine-learning models that reasonably distinguish between three different categories of emotions. The contributions of our paper are three-fold:

- A presentation of results obtained from an experiment with 18 participants and the publication of the resulting data set[1].
- An investigation of the influence of five physical activities on physiologically-based emotion recognition.
- New features to recognize emotions interfered by physical activities.

The rest of the paper is organized as follows: In Section II, we present the state of the art of emotion recognition, focusing on systems that utilize physiological responses to emotions and to stress. In Section III, we derive research questions based on the related work, setting the aim of this paper. In Section IV, we outline the underlying emotion model as well as emotion categories and describe the setup of our experiment. In Section V and VI, we detail our approach, elaborating on pre-processing and features that facilitate emotion recognition during physical activities. Finally, we discuss the results of this research.

## II. Related Work

The number of wearables already embedding physiological sensors is continuously rising [9]. On the one hand, the pervasiveness of such devices increases the amount of physiological data, covering various facets of our everyday life. Also the fact that physiological signals cannot be easily suppressed and

[1] https://www.comtec.eecs.uni-kassel.de/emotiondata/

978-1-5386-9151-9/19/$31.00 ©2019 IEEE

controlled by individuals, compared to emotion recognition via gestures or facial expressions [6], increase the research interest in those wearables. On the other hand, physiological sensors introduce new challenges to the field of emotion recognition. For example, challenges include environmental influences such as ambient temperature changes, physical activities, or the consumption of caffeine, sugar, and other non-emotional factors [10].

Previous research has already investigated influences of physiological sensors in the field of emotion recognition [10]–[12]. For example, Picard et al. found that physiological signals of one person vary from day-to-day [10]. Furthermore, they found that this day dependence could be handled by applying Sequential Floating Forward Search followed by Fisher Projection. This method led to an accuracy of 81% for classifying eight emotions of one participant over 20 different days. Xu et al. investigated the after-effects of physical activities on emotion recognition with physiological signals [11]. Classification accuracies of approximately 20% were achieved with models trained on unaffected data sets when testing on data containing after-effects. To improve the overall classification accuracy, Heinisch et al. merged the aforementioned data sets and applied a selection of commonly used features for emotion recognition [12]. They achieved classification accuracies of up to 96%. The influence of physical activities on physiological signals has also been investigated in the field of stress detection [7], [8], [13]. In [13], Alamudun et al. studied the subject dependence and the influence of activities on stress recognition. By leaving one activity for each participant out, they reached a mean classification accuracy of 66% over 14 participants and four activities. Hong et al. found an accuracy decrease of 14% training with physically non-interfered stress data and testing with data influenced by strenuous activities [7]. To investigate physical responses to stressors in multiple stimuli scenarios, Hong et al. proposed the use of a two-stage classification for stress recognition. Based on the classified activity (first-stage), a corresponding stress recognition model was applied (second-stage). They achieved a mean classification accuracy of around 82% over 19 participants. Ramos et al. improved the two-stage classification proposed in [7] to handle the influence of physical activities on stress detection [8]. They modified the first-stage by introducing a clustering algorithm. They further trained activity independent models with the clustered data. With this approach, they achieved an accuracy of 65%, which was lower than the two-stage method of [7].

Motivated by these approaches, this paper address the influences of physical activities on physiologically-based emotion recognition. As there is a vast amount of physical activities we might perform during the day, emotion recognition models independent to activities still remain an issue. The influence of physical activities on stress detection has been already successful addressed by [7], [8]. However, there is still a dependency on physical activities, caused by the creation of separate stress detection models in the second stage. This might increase the overall effort and complexity of classification models. The same effect is involved in training an emotion classification model with emotion data influenced by a range of different physical activities. In the light of the results of Picard et al. [10], there are still open questions about the significance and generality of different features on emotion recognition.

## III. Goals & Hypotheses

A human being's physiological signals are influenced through many factors such as the environment or physical activities [10]. For robust and efficient emotion recognition, models have to cope with interferences caused, for example, by physical activities. We refer to the term *non-sensitivity* when pointing towards the ability to cope with interferences of physical activities. Motivated by existing approaches and studies that already focus on physiologically-based emotion recognition, we stress the following research questions:

- Can emotion recognition models be trained to be non-sensitive to physiological interferences (*RQ1*).
- Are non-sensitive emotion recognition models robust or are they subject-dependent and susceptible to segmentation parameters (*RQ2*).

In the next section, we present the underlying emotion model and detail the setup and scenarios of our experiment.

## IV. Emotion Model & Experimental Setup

In our experiment, we used the emotion model by Mehrabian and Russel to categorize emotions [14]. This three-dimensional model classifies emotions in the dimensions of pleasure, arousal, and dominance. Furthermore, we used the International Affective Digital Sounds System (IADS) [15] to elicit emotions while performing physical activities. Among others, this system contains sounds that relate to the following emotion categories, see Table I.

TABLE I
SELECTED EMOTION CATEGORIES AND THEIR SELF-ASSESSMENT MANIKIN SCALE RATING

| | |
|---|---|
| **High Positive Pleasure High Arousal (*HPHA*)** | |
| pleasure: $6.06 - 7.9$ | arousal: $6 - 7.54$ |
| **High Negative Pleasure High Arousal (*HNHA*)** | |
| pleasure: $1.57 - 2.92$ | arousal: $6.07 - 8.16$ |
| **Neutral (*NEUTRAL*)** | |
| pleasure: $4.18 - 5.64$ | arousal: $4.6 - 5.48$ |

The numbers that are given for each category refer to the rating of sound samples in Self-Assessment Manikin-Scale [16]. Physiological measurements were recorded using the biosignalsplux toolkit [17] and an E4-wristband [18]. Furthermore, we employed smartphone embedded acceleration, gyroscope, gravity, and orientation sensors to record data about a participant's physical activities: sitting, standing, walking, ascending and descending stairs. We placed the smartphone inside a participant's pocket. We used the same locations for the physiological sensors of the biosignalsplux toolkit as in our previous study [12]. The E4-wristband was located on the non-dominant hand and was used to gather a participant's Skin Temperature (ST), the movement with a three-axis acceleration sensor, the Electrodermal Activity (EDA), and the Blood Volume Pulse (BVP). We recorded data from 21 healthy participants - 11 females and 10 males, between 19 and 50 years of age. The data of three participants were omitted due to erroneous and missing physiological signals, resulting in $\approx 300$ minutes of physiological data in total.

### A. Scenarios

To reduce potential bias, we divided the participants into two groups. Participants from both groups started with the Scenario Activity (*S-A*) continuing either with the Scenario Emotion (*S-E*) or the Scenario Emotion with Activity (*S-EA*) before completing the study with the remaining scenario, respectively. Each participant was measured individually. Fig. 1 details the procedure of the considered scenarios.

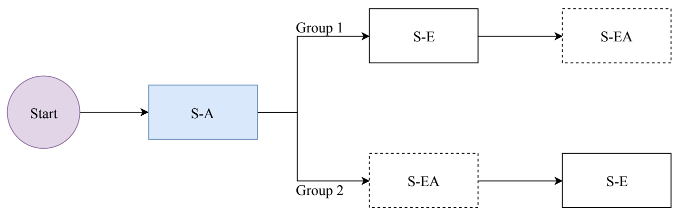



Fig. 1. Scenario procedures for different groups of participants: Scenario Activity (S-A), Scenario Emotion (S-E), Scenario Emotion with Activity (S-EA)

*Scenario Activity (S-A):* The participant was asked to perform physical activities without any elicitation of emotions. The scenario started with three minutes of resting. After that, the participant performed physical activities, i.e., sitting, standing, walking, ascending and descending stairs, each for a period of approximately 20 seconds.

*Scenario Emotion (S-E):* In this scenario, the participant sat in a quiet environment, listening to the sounds of each emotion category via headphones to prevent environmental interferences. For each considered emotion category, we chose sound samples for a total period of 2 minutes. We started with NEUTRAL sound samples. Then, we played the sounds of the HPHA category. To neutralize the influence of the HPHA sounds, we played NEUTRAL sound samples again, before playing HNHA sounds. Finally, NEUTRAL sound samples were played again.

*Scenario Emotion with Activity (S-EA):* Finally, we combined both scenarios where emotions were elicited *while* a participant was performing physical activities. Each participant was asked to perform physical activities in the same order and time as in scenario *S-A* without resting but while listening to the sounds of one emotion category for each trail. The emotions were the same as in scenario *S-E*: first NEUTRAL, followed by HPHA and finally HNHA. After each trial of the full set of physical activities, the participant was sitting on a chair and listening to the sound samples related to NEUTRAL again to neutralize the participants' emotional state.

## V. Methodology

In this section, we present the steps towards emotion recognition models, non-sensitive to physical activities. We elaborate on preprocessing and filtering techniques as well as describe the features used in our evaluation.

### A. Data Pre-processing

Different kinds of noise (e.g., caused by moving cables or gaps between the skin and the electrodes) were observed in the biosignalsplux sensor data. To reduce the noise, we applied several filtering techniques for each physiological sensor, shown in Table II. The Electromyogram signal (EMG) was filtered in two different ways. First, we filtered the signal with a fifth-order high-pass Butterworth filter with a cut-off frequency of $40Hz$ - EMG (H). Second, we used a fourth-order low-pass Butterworth filter with a cut-off frequency of $5Hz$ on the raw signal - EMG (L). Furthermore, a fourth-order low-pass Butterworth filter with a cut-off frequency of $0.5Hz$ and $0.25Hz$ was used to filter the EDA and ST signals, respectively. Before we filtered the Piezoelectric Respiration signal (PZT), a roll median function was used. Then, we filtered the PZT signal with a first-order low-pass Butterworth filter and a cut-off frequency of $1Hz$. Finally, the PZT signal was normalized. We decided not to filter the E4-wristband signal, as no significant noise was observed.

TABLE II
Filtering Techniques applied on biosignalsplux data

| Sensor | Filtering | Units |
|---|---|---|
| EMG (H) | High-pass filter (40Hz, 5th order) | Micro Volt |
| EMG (L) | Low-pass filter (5Hz, 4th order) | Micro Volt |
| EDA | Low-pass filter (0.5Hz, 4th order) | Micro Siemens |
| ST | Low-pass filter (0.25Hz, 4th order) | Celsius |
| PZT | Rollmedian (7 values, extend),<br>Low-pass filter (1Hz, 1st order),<br>Normalization | Percentage |

### B. Window Size

To assess the robustness of emotion recognition models, non-sensitive to physical activities, we also wanted to investigate the influence of the segmentation parameters, especially of the window size. For our analysis, we used the sliding window algorithm to segment our sensor data and analyzed the influence of different window lengths. We increased the window lengths from $100ms$ to $600ms$ in $50ms$ steps and evaluated the data.

### C. Features

For our evaluation, we used 15 statistical features on each physiological signal (e.g., mean, standard deviation or the mean of the absolute value of the first difference) [19]. As we have seen in our last paper, the slope of the linear regression line was able to distinguish between different emotion categories for ST [12]. Therefore, we further investigated features based on the linear regression line in this paper.

For calculating the linear regression line, we used SciPy, an open-source mathematics library for Python [20]. The $linregress$ function takes two datapoints and calculates a linear least-squares regression. For our evaluation, we further processed the slope of the regression line, as well as its intercept. Let $W = (x_1, x_2, ...x_n)$ be a window with length of $n$ and $I = (1, 2, ...n)$ the corresponding $index$ of the elements in $W$. The features are then defined as

$$f_{slope} = \sqrt{|slope(linregress(I, W))|} \quad (1)$$

$$f_{1intercept} = \sqrt{|intercept(linregress(I, W)|} \quad (2)$$

$$f_{2intercept} = \sqrt{|intercept(linregress(I, W)|}^3 \quad (3)$$

After a preliminary analysis of the signals, we found, that some features and sensors were more relevant for the classification than others. Therefore, we evaluated the performance of the models with a second set of features, namely: the mean of the absolute values of the first differences, the absolute value of the slope of the linear regression line, the square root of the absolute value of the intercept of the linear regression line, and the third power of the square root of the absolute value of the intercept of the linear regression line. These selected features were calculated on the BVP of the E4-wristband, the ST of the E4-wristband, the EMG (H) and the EMG (L) signals of the biosignalsplux toolkit. In our analysis, we found that the ST, the EMG, and the BVP were useful for classifying the three emotion categories.

## VI. Evaluation

This section describes and compares the results of the evaluation. To investigate the first research question, we trained our models with the physiological signals influenced by the least strenuous activity (data set *S-E*) and tested with data influenced by more strenuous activities (*S-EA*). Then, we separated the data by activities from the data set of scenario *S-EA* and used the data of each activity in the testing phase to evaluate our classifiers empirically. Finally, we investigated the impacts of different window lengths on the classification performance. For the classification, we chose the three best classifiers from our previous research, namely Decision Tree (DT), Random Forest (RF) and K-Nearest Neighbor (KNN, with k=3) [12]. For each participant, the classification was done 10 times for all classifiers to reduce the bias.

Fig. 2 depicts the mean classification accuracy of all participants and for all activities of *S-EA*. The KNN classifier achieved the best accuracy over all window sizes. Rather than using all features, where the mean classification accuracy is ranging from 35% - 50% for all classifiers, the set containing only a selection of features yielded a higher classification accuracy ranging from 56% - 67%. Also, the DT, as well as the RF classifier, achieved similar results.

In addition to the classification performance, we were interested in the performance considering single emotion categories. Fig. 3 shows the mean f-measure of all participants and window sizes for the selected feature set. We observe that all emotion categories were fairly recognized by the classifiers, with mean f-measures ranging from 0.49 - 0.79. In particular, we note that the classifiers achieved higher performances recognizing the *NEUTRAL* and *HNHA* emotion categories. Further analysis showed that high arousal categories of emotions were more difficult to distinguish. In case of

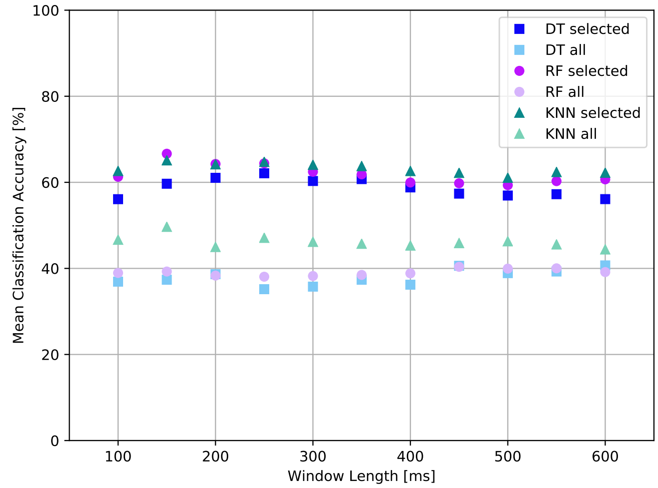



Fig. 2. Mean classification accuracy of all participants with different feature sets as training data

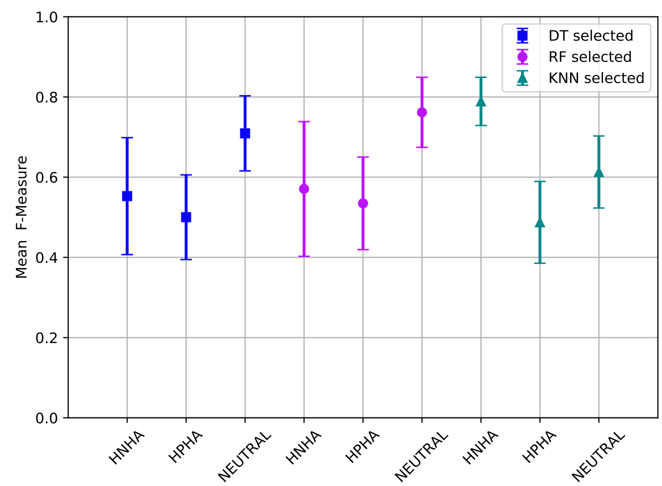
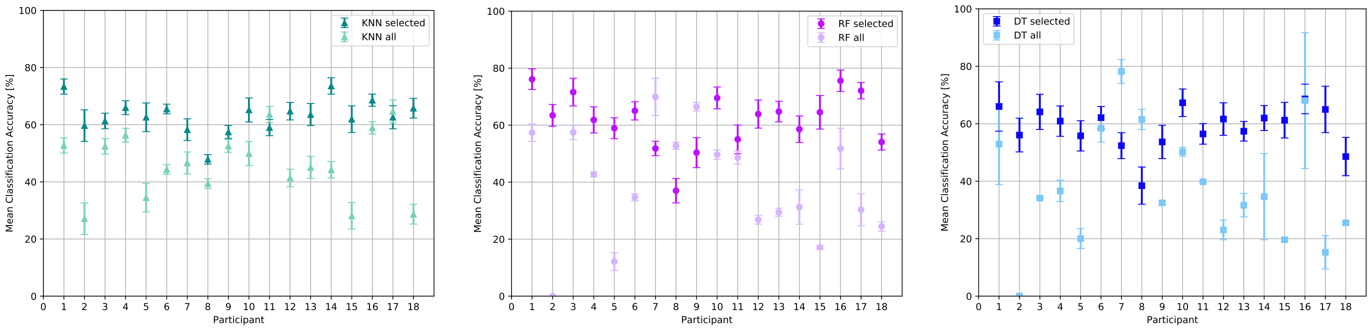



Fig. 3. Mean f-measure over all participants and window sizes using the data of the selected feature set as training data

miss-classification, we observed that the *HPHA* category was incorrectly classified as *NEUTRAL*. However, this occurred less often than the miss-classification with *HNHA*. Also, *NEUTRAL* was rarely miss-classified as *HPHA* or *HNHA*. Consequently, the f-measures of *HPHA*, tend to be lower than *NEUTRAL* and *HNHA*.

Considering the second research question, we investigated the impact of different window sizes on the classification performance. Fig. 4 depicts the mean classification accuracy and standard deviation over all windows for each participant. We observe that the standard deviations are different for each classifier and participant. The KNN shows the lowest standard deviation for all participants, followed by the RF. Also, we note that the all classifiers performed better on the data of the selected feature set than on the data of all features.

We also extracted the data of all single activities from the data set of scenario *S-EA* to investigate the accuracy of our




(a) K-Nearest Neighbors (KNN) (b) Random Forest (RF) (c) Decision Tree (DT)

Fig. 4. Mean classification accuracy and standard deviation over all windows for each participant

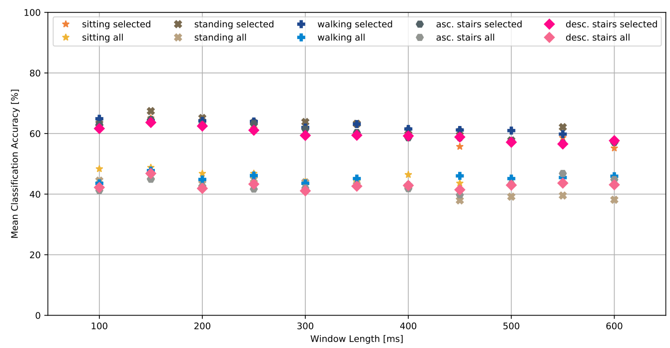



Fig. 5. Mean classification accuracy over all participants for single activities using K-Nearest Neighbors (KNN)

classification models for each activity in the testing phase individually. Note that the classification models were only trained with data from scenario *S-E* influenced by a low strenuous activity. Fig. 5 shows the mean classification accuracy of all participants and for single activities using the KNN classifier. Analog to the aforementioned results, we note that the models trained with the data of the selected features achieved higher classification accuracies than the models trained with the data of all features. Using the selected feature set as training data, emotions were recognized for all activities with a classification accuracy ranging from $55\%$ to $67\%$. For the RF and DT we found similar results ranging from $54\%$ to $71\%$, and from $51\%$ to $65\%$, respectively.

## VII. Discussion

Considering the research questions, we were first interested whether machine-learning models could be trained independently to physiological interferences caused by physical activities (*RQ1*). To answer this question, we chose to train emotion models on physiological data influenced by the least strenuous activity (*S-E*) and tested the performance against the remaining, more strenuous activities (*S-EA*). Overall, our results indicate that the three emotion categories, *NEUTRAL*, *HPHA*, and *HNHA*, can be recognized, ranging from $56\% - 67\%$ classification accuracy for selected feature sets over all window sizes and per participant. The *NEUTRAL* category achieved the highest f-measure followed by *HNHA* using the RF and DT classifier. An exception was the KNN classifier which achieved higher classification accuracies on the *HNHA* category than the *NEUTRAL* category. However, we noticed that the high arousal emotions, i.e., *HPHA*, *HNHA* were confused with another for all participants. The reason for this might be that the features corresponding to emotion categories of being high arousal are similar in their physiological signal responses. *NEUTRAL* is more often confused with *HPHA* than with *HNHA*. A conceivable cause might be the consequences of selecting the sound samples for the *HPHA* and *NEUTRAL* categories, which are closer together on the pleasure scale than to *HNHA*. This decision was made to have two minutes of sounds available for each emotion category. Regarding the robustness of non-sensitive emotion recognition models (*RQ2*), we noticed that the size of the sliding window did not have a significant effect on the classification accuracy, in all considered cases. Considering the selected feature, standard deviations of the classification accuracy range over all participants from $1.64\% - 5.1\%$ for the KNN and $2.52\% - 5.89\%$ for RF, and $3.43\% - 8.59\%$ for DT classifier. Nonetheless, we note some outliers where the window size has a great impact on the classification accuracy, expressed as a high standard deviation. For example, a significant impact of different window sizes on the classification accuracy was observed for the DT using as training set the data of all features with a standard deviation of $23.62\%$ for participant 16. We assume that for these outliers some window lengths contain more information useful to the classifiers to distinguish between emotion categories. The reason for this might be that emotions and the influence of physical activities on physiological signals are subject dependent due to personal characteristics or the individuality in the execution of physical activities.

## VIII. Conclusion

In this paper, we investigated emotion recognition models non-sensitive to interferences of physical activities. In particular, we considered the influence of five physical activities namely sitting, standing, ascending and descending stairs on physiologically-based emotion recognition. To evaluate non-sensitive models, we trained three different classification algorithms with the data of the least strenuous physical activity (sitting) and tested with the data of the remaining, more strenuous activities. Using the data of selected features, partly based on the linear regression line, we achieved mean classification accuracies between $55\%$ and $67\%$ for classifying three different emotion categories. Classification accuracies between $48\%$ and $73\%$ were achieved when considering individual activities and participants. We found that the data of features based on the linear regression line improved the performance of emotion recognition models. The relative enhancement was approximately $+20\%$ over using the data of all features. For this new set of features, we found no significant influence of different window lengths on the classification performance in all considered cases.

Our results show that the complexity of emotion classification models can be minimized by applying the proposed features and classification algorithms. Contrary to already existing approaches, our approach only requires one model to distinguish three emotion categories interfered by physical activities. However, some research questions remain. In future work, we plan to investigate more strenuous activities as well as new features and sensor combinations to provide more reliable information about emotional states “in-the-wild”.

## Acknowledgment

The authors would like to thank all the participants of the experiment for their support.

## References

[1] M. Raja and S. Sigg, “RFexpress! - RF Emotion Recognition in the Wild,” in *2017 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)*. Hawaii, HI, USA: IEEE, March 2017, pp. 38–41.

[2] T. Fritz and S. C. Müller, “Leveraging biometric data to boost software developer productivity,” in *2016 IEEE 23rd International Conference on Software Analysis, Evolution, and Reengineering (SANER)*, vol. 5. Suita, Japan: IEEE, March 2016, pp. 66–77.

[3] 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, pp. 1–41, Jun. 2018.

[4] S. Jerritta, M. Murugappan, R. Nagarajan, and K. Wan, “Physiological signals based human emotion recognition: A review,” in *2011 IEEE 7th International Colloquium on Signal Processing and its Applications*, Penang, Malaysia, March 2011, pp. 410–415.

[5] K. H. Kim, S. W. Bang, and S. R. Kim, “Emotion recognition system using short-term monitoring of physiological signals,” *Medical and Biological Engineering and Computing*, vol. 42, no. 3, pp. 419–427, May 2004.

[6] H. Gunes and M. Pantic, “Automatic, Dimensional and Continuous Emotion Recognition,” *International Journal of Synthetic Emotions (IJSE)*, vol. 1, no. 1, pp. 68–99, Jan. 2010.

[7] J.-H. Hong, J. Ramos, and A. K. Dey, “Understanding physiological responses to stressors during physical activity,” in *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*. New York, NY, USA: ACM, September 2012, pp. 270–279.

[8] J. Ramos, J.-H. Hong, and A. K. Dey, “Stress recognition - a step outside the lab,” in *Proceedings of the International Conference on Physiological Computing Systems - Volume 1: PhyCS*, vol. 1, INSTICC. Lisbon, Portugal: SciTePress - Science and and Technology Publications, 2014, pp. 107–118.

[9] “Global Wearables Market Grows 7.7% in 4Q17 and 10.3% in 2017 as Apple Seizes the Leader Position, Says IDC.” [Online]. Available: https://www.idc.com/getdoc.jsp?containerId=prUS43598218

[10] R. W. Picard, E. Vyzas, and J. Healey, “Toward machine emotional intelligence: analysis of affective physiological state,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 10, pp. 1175–1191, Oct. 2001.

[11] Y. Xu, I. Hübener, A.-K. Seipp, S. Ohly, and K. David, “From the Lab to the Real-world: an Investigation on the Influence of Human Movement on Emotion Recognition using Physiological Signals,” in *2017 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)*, Kona, HI, USA, 2017, pp. 345–350.

[12] J. S. Heinisch, I. Hübener, and K. David, “The Impact of Physical Activities on the Physiological Response to Emotions,” in *2018 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)*, Athens, Greece, 2018, pp. 824–829.

[13] F. Alamudun, J. Choi, R. Gutierrez-Osuna, H. Kahn, and B. Ahmed, “Removal of Subject-Dependent and Activity-Dependent Variation in Physiological Measures of Stress,” in *Proceedings of the 6th International Conference on Pervasive Computing Technologies for Healthcare*, San Diego, CA, USA, March 2012, pp. 115–122.

[14] A. Mehrabian and J. A. Russel, *An Approach to Environmental Psychology*. Cambridge, MA, USA: the MIT Press, March 1974.

[15] M. M. Bradley and P. J. Lang, “The international affective digitized sounds (2nd edition; iads-2): Affective ratings of sounds and instruction manual,” University of Florida, Gainesville, FL, USA, Tech. Rep. B-3, 2007.

[16] ——, “Measuring emotion: The self-assessment manikin and the semantic differential,” *Journal of behavior therapy and experimental psychiatry*, vol. 25, no. 1, pp. 49–59, 1994.

[17] “biosignalplux — wearable body sensing platform.” [Online]. Available: http://biosignalsplux.com/en/

[18] M. Garbarino, M. Lai, D. Bender, R. W. Picard, and S. Tognetti, “Empatica e3 - a wearable wireless multi-sensor device for real-time computerized biofeedback and data acquisition,” in *2014 4th International Conference on Wireless Mobile Communication and Healthcare - Transforming Healthcare Through Innovations in Mobile and Wireless Technologies (MOBIHEALTH)*, Athens, Greece, Nov. 2014, pp. 39–42.

[19] E. Vyzas and R. W. Picard, “Affective pattern classification,” The Media Laboratory Massachusetts Institute of Technology, Tech. Rep. FS-98-03, 1998.

[20] E. Jones, T. Oliphant, P. Peterson *et al.*, “Scipy: Open source scientific tools for python,” 2001–. [Online]. Available: http://www.scipy.org/