Emotion Recognition of Physical Activities for Health Monitoring Using Machine Learning

Sun, Kai; Zhu, Jihong; Liang, Jie

DOI
10.1109/APCCAS55924.2022.10090279

Publication date
2022

Document Version
Final published version

Published in
Proceedings of the IEEE Asia Pacific Conference on Circuits and Systems, APCCAS 2022

Citation (APA)
Sun, K., Zhu, J., & Liang, J. (2022). Emotion Recognition of Physical Activities for Health Monitoring Using Machine Learning. In *Proceedings of the IEEE Asia Pacific Conference on Circuits and Systems, APCCAS 2022* (pp. 400-403). IEEE. https://doi.org/10.1109/APCCAS55924.2022.10090279

The 2022 IEEE Asia Pacific Conference on Circuits and Systems

Emotion Recognition of Physical Activities for Health Monitoring Using Machine Learning

Kai Sun
School of Microelectronics
Shanghai University
Shanghai, China
sunkai1997@shu.edu.cn

Jihong Zhu
Cognitive Robotics
TU Delft
Delft, Netherlands
j.zhu-3@tudelft.nl

Jie Liang*
School of Microelectronics
Shanghai University
Shanghai, China
liangjieclair@shu.edu.cn

2022 IEEE Asia Pacific Conference on Circuits and Systems (APCCAS) | 978-1-6654-5073-7/22/$31.00 ©2022 IEEE | DOI: 10.1109/APCCAS55924.2022.10090279

***Abstract*—Emotion recognition based on physiological data has attracted increasing attention in physiological monitoring, affective computing, and other fields. This paper proposes a method to classify human's emotion for health monitoring in physical activities by using machine learning. Participants completed the experiment including walking, running, and other physical activities. The data of photoplethysmography (PPG) and electrodermal activity (EDA) were recorded by wearable sensors on participants. After the data processing and feature extraction, two classifiers, support vector machine (SVM) and random forest (RF) were applied independently on the dataset to classify human's emotion, including calm, excited, relaxed, bored, and afraid. As a result, the SVM classifier achieved an accuracy of 81.87% and the accuracy of RF classifier is 86.61%. These results demonstrated the effectiveness of the proposed method on emotion recognition in human's physical activities.**

Keywords—wearable sensor, electrodermal activity, machine learning, physical activities

I. Introduction

The change of one's emotion will normally bring changes on one's expression, behavior, and physiological parameters. Emotion study has been attracting interest by researchers during the past decades. Russell [1] first proposed a two-dimensional model of valence-arousal to measure feelings in 1980. In this theory, valence represents the positive and negative of emotion, and arousal represents the intensity of emotion, which is a measurement from calmness to excitement. Physiological data, such as electrodermal activity (EDA) signal and photoplethysmography (PPG) signal, show changes of people's emotion truer and more objective than sensory expression like expression and voice, which controlled by people's subjective consciousness easily. Emotion recognition that classify different kinds of emotions by collecting and analyzing human's various physiological data, has driven important attention in physiological monitoring, affective computing, and other fields, owing to the rapid development on interdisciplinary study of wearable biosensors and machine learning.

However, most of the recent research on emotion recognition focuses on low-intensity activities such as sleep, daily work, etc. Aasim Raheel et al. [2] used physiological data, electroencephalography (EEG) , PPG and galvanic skin response (GSR) observed in response to tactile enhanced multimedia content, classifying four (happy, relaxed, angry, and sad) emotions with K nearest neighbor (KNN) classifier. Al Machot et al. [3] used EDA signals based on deep-learning to classify human emotions, with the accuracy for subject-independent classification of 85%. In fact, the influence of psychological states is as important as physical exertion to human's physiological rhythms and health that should not be underestimated in physical activities. For example, EDA acts as an indicator for sympathetic activation due to the stress reaction while the sweat glands are also controlled by the sympathetic nervous system (SNS) [4]. And heart rate variability of PPG reflects the autonomic balance between the SNS and parasympathetic nervous system (PNS), the aerobic-based adaptation and fatigue status of humans are also could be read from PPG signal [5]. In physical activities, emotion changes associated with SNS and PNS could reveal the psychological states and physiological status of our bodies for health monitoring.

Therefore, this paper proposes a method to recognize human's emotion for healthy psychological states in physical activities by using machine learning. EDA and PPG signals were collected while participants took physical activities with a range of intensity from low to high by wearable sensors on their fingers. After data pre-processing works like filtering and wavelet transforming, we can extract features from the collected data to generate dataset for classifying human emotions using machine learning in time-domain and frequency-domain. Two classifiers, support vector machine (SVM) and random forest (RF), were applied independently on the dataset to predict human's five emotions, including calm, excited, relaxed, bored, and afraid. Our results show that the SVM classifier achieved an accuracy of 81.87% and the accuracy of RF classifier is 86.61%, it demonstrated the effectiveness of this proposed method using in human's sport and physical activities. Overview of the emotion recognition method is illustrated in Fig. 1.

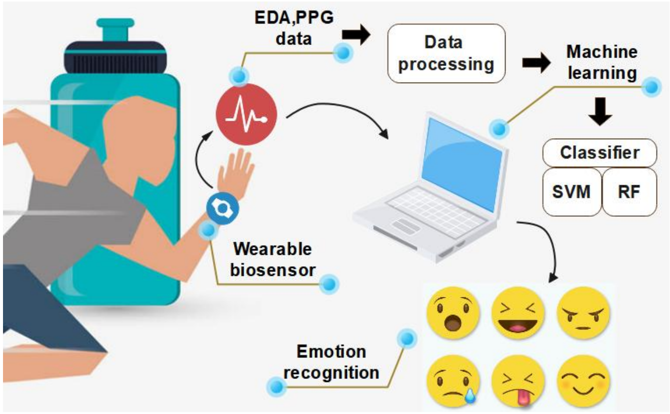

Fig. 1. Overview of the emotion recognition method.

II. Data Collection

A. Participants

Six young and healthy participants (5 males and 1 female, between 22 and 30 years old), participated in this study voluntarily. The experimental study was designed according to the Helsinki declaration and approved by School of Microelectronics, Shanghai University.

B. Procedure

The physiological data acquisition device consists of a microcontroller (ESP-WROOM-32, Espressif, Inc., China) for controlling and processing the system, GSR sensor (Grove v2.0, Sichiray, Inc., China) and PPG sensor (MAX30102, Jiaqisheng, Inc., China) were used for collecting EDA signal with a sampling rate of 40 Hz and PPG signal with 50 Hz. The placement and devices of experiment are shown in Fig. 2.

Participants were asked to take physical activities described below, and sensors were attached to their fingers.

- 90-second-long baseline (sitting on a chair)
- 150-second-long low-intensity physical activity (walking with speed of 4 km/h)
- 150-second-long moderate-intensity physical activity (running with speed of 8 km/h)
- 120-second-long high-intensity physical activity (take squats at least 30 times)
- 90-second-long high-intensity physical activity (jumping jack uninterruptedly)

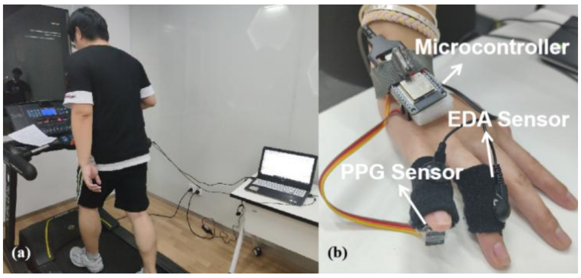

Fig. 2. The experimental procedure. (a) Placement of the experiment. (b) The structure of physiological data acquisition devices.

After each physical activities, participants should have a rest for 5 minutes and take questionnaires of Borg's Rating of Perceived Exertion (RPE) and Self-Assessment Manikin (SAM) to rate how they just felt from -4 to 4. RPE scale contains three physical activity relative intensity classes, namely the sedentary, moderate, and vigorous. SAM is a graphical and verbal tool to measure the user's affective reaction in response to a variety of stimuli in terms of arousal which can be measured by RPE scale and valence [6]. Five emotions (shown in Fig. 3), calm, excited, relaxed, bored, and afraid, were categorized based on different values of valence and arousal. For example, if (2, 3) is the answer of one's SAM scale about valence and arousal, his emotion will be targeted as excited in the physical activity.

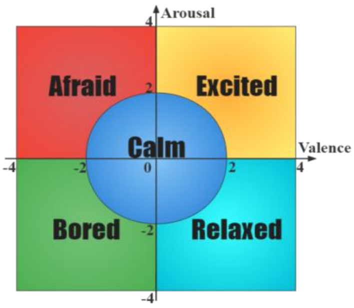

Fig. 3. Classification of five emotions in SAM scale.

III. Data Processing

We generate the dataset from raw sensor data for machine learning. The data processing of the emotion recognition method consists of three steps: pre-processing, feature extraction, normalization and feature selection.

A. Pre-processing

The data we collected from wearable sensors are analog signals susceptible to noises as 50 Hz power frequency and myoelectricity interference. Generally, the frequency of EDA and PPG signal are both in the low frequency range while the frequency of noises generated from human's movement and environment is in the range of high frequency. Butterworth filters (low-pass, frenquency-pass = 10 Hz, order = 4; high-pass, frenquency-pass = 1 Hz, order = 4) were designed to remove the noises and baseline drift successively. In processing PPG signals, we need to detect as many R peaks as possible to obtain RR intervals (the time elapsed between two successive R-waves) and other features after filtering. The EDA and PPG signals from the first participant's 90-second-long jumping jack activity before and after pre-procesing is shown in Fig. 4. Finally, using wavelet transform to adjust the resolution adaptively to analyse the signals in time and frequency domain from Equ. (1).

$$W(a,b)=\frac{1}{\sqrt{a}}\int_{-\infty}^{+\infty}f(t)\psi(\frac{t-b}{a})dt \tag{1}$$

After setting the dilation factor a, translation factor b, and wavelet type, mother wavelet $\psi(t)$ will transfer into wavelet basis function $W(a,b)$ about the function $f(t)$ we input with collected physiological data. The result of time-frequency analysis using wavelet transform of signals is shown in Fig. 5.

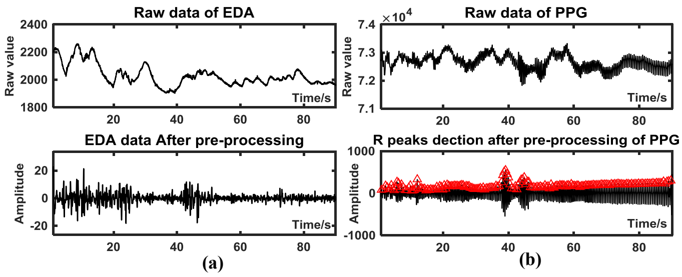

Fig. 4. Raw and processed signals of (a) EDA and (b) PPG.

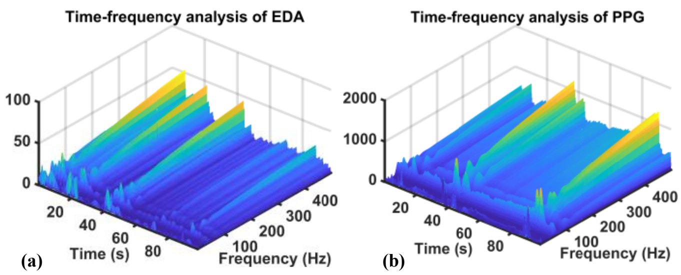

Fig. 5. Time-frequency analysis using wavelet transform of (a) EDA and (b) PPG.

B. Feature Extraction

A number of time and frequency domain features were extracted from each sensor modality after cutting the processed data into 1-second windows. We extracted five time-domain features (Mean, Max, Min, KURT, SKEW) from the collected EDA data. Four time-domain features (RRI, HR, SDNN, RMSSD) and two frequency-domain features (SD1, SD2) were extracted after detecting R peaks and computational analysis from the collected PPG data. These features are related to the valence and arousal of motion perception and emotional stimuli, form the dataset to help us classify human's emotion in sports and physical activities [7]. The extracted features from the sensor modalities are given in Table Ⅰ.

TABLE I. FEATURE SET EXTRACTED FROM EACH SENSOR

Physiological Signal	**Feature**	**Description**
EDA	Mean	Mean of EDA
	Max	Maximum value of EDA
	Min	Minimum value of EDA
	KURT	Kurtosis of EDA
	SKEW	Skewness of EDA
PPG	RRI	RR-interval, the time elapsed between two successive R-waves
	HR	Heart Rate
	SDNN	Standard deviation of Normal RR intervals
	RMSSD	Root mean square of successive RR interval differences
	SD1	Standard deviation perpendicular the line of identity
	SD2	Standard deviation along the line of identity

C. Normalization and Feature Selection

Since the baseline of EDA data varied greatly among participants, and the range of characteristic values in 11 extracted statistical features is in different orders of magnitude. In order to limit features to a common range and optimize the quality of input data, linear methods were used to normalize each feature to a zero mean and unit variance between (0, 1). For every participant, 600 groups of normalized data including 11 features and one target of emotion were extracted from 600-second recorded data. Due to occasional fluctuations in data collection by experimental devices, few anomalous data were eliminated and blank data were filled in with the average value of the four surrounding data. Finally, the generated dataset contains 3375 instances with 37125 features was fed into the classifiers as inputs.

IV. METHODOLOGY

A. Classification Algorithms

To solve the nonlinear and multi-classification problem of five emotions, two widely used machine learning algorithms, support vector machine (SVM) and random forest (RF), were applied independently on the dataset to classify human's emotion in our work.

SVM is a binary classification model which is defined in feature space with the largest interval. It uses nonlinear kernel function to map the input data into a high-dimensional space for linear classification, so as to construct an optimal hyperplane [8], that is promising in solving nonlinear multi-classification problem. According to the distribution of the physiological data in the dataset, the RBF kernel that has the advantage of fast convergence was selected as the kernel function to improve the SVM classifier performance.

Random forest (RF) is an intelligent algorithm combining decision tree (DT) and bootstrap resampling method, which has good performance based on the information theory. Similar to the distribution of trees in a forest, many decision tree classifiers are constructed at work, and each decision tree is trained by information entropy independently based on different features and subsets input. At the time of testing, the optimal parameter and solution of the tree with most votes were given to be predicted results of RF classifier about human's emotion classification in physical activities.

Both classification algorithms SVM and RF were written in Python and verified by 25% set aside method and 5-fold cross-validation scheme.

B. Performance Evaluation

From the results given of the two classification algorithms, their performance could be evaluated based on classification of four predicted results including true positive (TP), true negative (TN), false negative (FN), and false positive (FP). The performance metrics of the emotion classification method are listed as equations below:

$$precision = \frac{TP}{TP + FP} \times 100\% \tag{2}$$

$$recall = \frac{TP}{TP + FN} \times 100\% \tag{3}$$

$$accuracy = \frac{TP + TN}{TP + TN + FN + FP} \times 100\% \tag{4}$$

$$F1Score = \frac{2 \times precision \times recall}{precision + recall} \times 100\% \tag{5}$$

The performance results of emotion recognition in our work are presented in Table Ⅱ. It shows that the SVM classifier achieved an overall accuracy of 81.87%, while the overall accuracy of RF is 86.61%, performs better than SVM.

TABLE II. CLASSIFICATION PERFORMANCE RESULTS OF EMOTION RECOGNITION USING SVM AND RF

Classi fier	**Emotion**	**Precision**	**Recall**	**F1 Score**	**Accuracy**
SVM	Calm	0.92	0.74	0.82	81.87%
	Excited	0.76	0.89	0.82	
	Relaxed	0.74	0.76	0.75	
	Bored	0.85	0.87	0.86	
	Afraid	0.85	0.83	0.84	
	Overall	0.82	0.82	0.82	
RF	Calm	0.93	0.75	0.83	86.61%
	Excited	0.87	0.91	0.89	
	Relaxed	0.81	0.90	0.85	
	Bored	0.86	0.85	0.85	
	Afraid	0.89	0.91	0.90	
	Overall	0.87	0.86	0.86	

The number of correctly and incorrectly classified results of two classifiers is shown in the confusion matrix (Fig. 6). The numbers in dark blue cells along the main diagonal represent the quantity of correctly classified results for five emotions (TP+TN), i.e., the number of predicted instances matches the true instances. While the numbers in light cells represent the quantity of incorrectly ones (FN+FP). The SVM classifier predicted 691 instances correctly in 844 test instances (25% of the total 3375 instances, the other 75% using for training), while the number of RF classifier is 731.

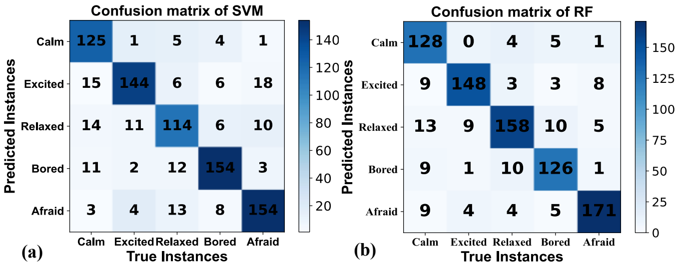

Fig. 6. Confusion matrix for (a) SVM classifier and (b) RF classifier, applied on collected physiological data of EDA and PPG.

From ROC (Receiver operating characteristic) curves that plotted with TP rate against FP rate of two classifiers shown in Fig. 7, we notice that two ROC curves both are found near the upper left corner, which indicates AUC (area under ROC curve, between [0.5, 1]) achieved values of 0.97 and 0.98, closer to 1, means SVM and RF classifiers both obtained excellent performance in the work of emotion recognition.

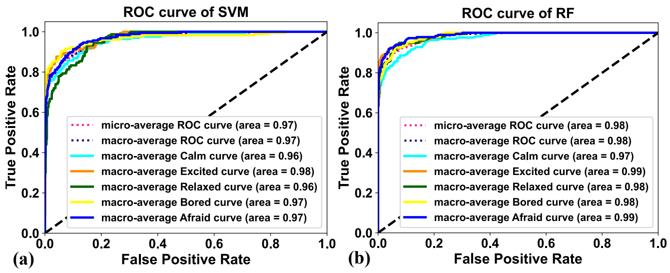

Fig. 7. ROC curve for (a) SVM classifier and (b) RF classifier, applied on collected physiological data of EDA and PPG.

Our proposed emotion recognition method achieved a higher accuracy while comparing with [2, 3, 9] as shown in Table III. Collected physiological data, classification of emotions and classifiers of algorithms in [2, 3, 9] are similar to our work. As can be seen that our work provide an efficient and promising method of emotion recognition to monitor psychological states for health in human's physical activities.

TABLE III. PERFORMANCE COMPARISON OF THE PROPOSED EMOTION RECOGNITION METHOD WITH RELATED METHODS

Method	Modality	Emotions	Classifier	Accuracy
[2]	EEG, GSR, PPG	Happy, Angry, Sad, Relaxed	KNN	79.76%
[3]	EEG, EDA, ECG	Happy, Angry, Sad, Relaxed	SVM	59.00%
			RF	74.00%
			CNN	85.00%
[9]	EEG	Happy, Anger, Sad, Love	SVM	75.62%
			ANN	78.11%
This work	EDA, PPG	Calm, Excited, Relaxed, Bored, Afraid	SVM	81.87%
			RF	86.61%

The classification of five kinds of emotions all perform well with similar parameters except for the emotion of relaxed, which has the smallest number of instances in the dataset, only 17.78% in the data of five emotions, may proved that moderate-high intensity physical activities bring people intense emotions rather than soothing ones like relaxation. Number in light cells indicates that emotions with similar physiological parameters as well as valence and arousal are easily incorrectly classified, such as excited and afraid, calm and relaxed, which are adjacent on the coordinate axis in SAM scale shown in Fig. 3. The performance of classifiers are also affected seriously by large fluctuations of EDA and PPG data collected. Movements in people's physical activities are bigger than they usually do in their daily life, thus, the collection of physiological data is so important that abnormal data with large fluctuations should be processed to obtain better performance of emotion classification in human's physical activities.

V. CONCLUSION

In the present work, we proposed a method to recognize emotions for humans healthy psychological states in physical activities using machine learning. SVM and RF classifiers were used to classify five emotions based on collected and processed physiological data of EDA and PPG from our experiment of physical activities. Two classifiers achieved accuracy of 81.87% and 86.61% respectively. The results show that our work using machine learning based on EDA and PPG physiological data is a theoretically-justified and efficient method to recognize emotions in human's physical activities. Further studies should increase the number of participants, improve the performance of classifiers, and combine with internet of things (IoT) and wearable biosensor technology to monitor psychological states and give advice for health timely by emotion recognition in human's sports and physical activities.

ACKNOWLEDGMENT

We would like to thank Dr. Jihong Zhu from TU Delft for his advice about the work. And thanks to participants from School of Microelectronics, Shanghai University.

REFERENCES

[1] Russell, James A. "A circumplex model of affect." *Journal of personality and social psychology* 39.6 (1980): 1161.

[2] Raheel, Aasim, et al. "Physiological sensors based emotion recognition while experiencing tactile enhanced multimedia." *Sensors* 20.14 (2020): 4037.

[3] Al Machot, Fadi, et al. "A deep-learning model for subject-independent human emotion recognition using electrodermal activity sensors." *Sensors* 19.7 (2019): 1659.

[4] Sequeira, Henrique, et al. "Electrical autonomic correlates of emotion." *International journal of psychophysiology* 71.1 (2009): 50-56.

[5] Schneider, Christoph, et al. "Heart rate monitoring in team sports—a conceptual framework for contextualizing heart rate measures for training and recovery prescription." *Frontiers in physiology* 9 (2018): 639.

[6] Bradley, Margaret M., and Peter J. Lang. "Measuring emotion: the self-assessment manikin and the semantic differential." *Journal of behavior therapy and experimental psychiatry* 25.1 (1994): 49-59.

[7] Poli, Angelica, et al. "Cross-Domain Classification of Physical Activity Intensity: An EDA-Based Approach Validated by Wrist-Measured Acceleration and Physiological Data." *Electronics* 10.17 (2021): 2159.

[8] Chapelle, Olivier, et al. "Choosing multiple parameters for support vector machines." *Machine learning* 46.1 (2002): 131-159.

[9] Bhatti, Adnan Mehmood, et al. "Human emotion recognition and analysis in response to audio music using brain signals." *Computers in Human Behavior* 65 (2016): 267-275.