Emotion recognition using physiological signals

Leonard Mirt

December 2023

Content

1 Introduction

Emotions are a basic component of our life, just like breathing or eating. They are also responsible for a majority of our decisions. For that reason, emotion recognition is an important and profitable research problem. The ability to recognize emotions can help in emotion-based disorders, autism, monitoring our well-being and mental health, controlling stress, human–computer interaction, recommendation systems, and computer games. At the same time, emotion recognition is a very ambitious task, as it connects several disciplines, i.e., psychology, electronics/sensors, signal processing, and machine learning.

The primary focus of this review is on emotion recognition from the physiological signals because it can be performed continuously in everyday life using wearables, as opposed to facial and speech emotion recognition.

A part of this paper explores the methods used in [1]Quiroz JC, Geangu E, Yong MH 2018 paper and the other tries to improve the results by using various methods.

2 Related Work

Quiroz JC, Geangu E, Yong MH 2018 paper[1] focuses on the exploration of movement sensor data from a smartwatch as a means to infer an individual's emotional state. The investigation involved a user study with 50 participants. Employing a mixed-design study approach—incorporating within-subjects factors (emotions: happy, sad, and neutral) and between-subjects factors (stimulus type: audiovisual "movie clips" and audio "music clips")—each participant experienced both emotions within a single stimulus type. Participants walked a distance of 250 meters while wearing a smartwatch on one wrist and a heart rate monitor strap on the chest. Additionally, they completed a short questionnaire (20 items; Positive Affect and Negative Affect Schedule, PANAS) before and after each emotional experience. Heart rate monitor data served as supplementary information. The analysis encompassed time series analysis on smartwatch data, t-tests on questionnaire items, and one-way analysis of variance on heart rate data. Features extracted from the time series were utilized to train and validate classifiers for emotion detection. Results revealed that participants, overall 50 young adults, reported feeling less negative affect after exposure to sad videos or sad music (P¡.006). In the realm of emotion recognition using classifiers, personal models outperformed personal baselines, achieving median accuracies exceeding 78

3 Methodology

3.1 Datasets

The dataset used in the first experiment is the same as in the original paper[1], heart-rate and movment data from a smart watch. The smart watch included a triaxial accelerometer and a triaxial gyroscope. The heart-rate was taken from heart rate monitor strap on the chest.

3.2 Data preparation

Walking times were categorized based on the emotional stimulus presented before each walking session. For instance, if a participant watched a movie clip known to induce happiness, all features extracted from the subsequent walking data were labeled as "happy." These labels were employed to train classifiers for distinguishing between happiness and sadness. Our findings include classifier results for the binary task of detecting happy versus sad emotions and the ternary task of detecting happy, sad, and neutral emotions.

Raw accelerometer data was initially filtered using a mean filter with a window size of 3. Sliding windows, consisting of one-second intervals (24 samples) with a 50

From each window of triaxial accelerometer and triaxial gyroscope data, we extracted 17 features, including mean, standard deviation, maximum, minimum, energy, kurtosis, skewness, root mean square, root sum square, sum, sum of absolute values, mean of absolute values, range, median, upper quartile, lower quartile, and median absolute deviation. These 17 features were extracted from each of the three axes of the accelerometer data and each of the three axes of the gyroscope data, resulting in a total of 102 features. Additionally, we calculated the angle between the signal mean and the x-axis, y-axis, and z-axis (3 features), standard deviation of signal magnitude (1 feature), and heart rate (1 feature), bringing the total features for a window's feature vector to 107. Unless specified otherwise, all 107 features were utilized for classification. However, we also explored classification performance based on features corresponding to specific sensors: accelerometer, gyroscope, and heart rate; accelerometer and heart rate; and accelerometer.

Data was segmented by condition, and individual models were constructed using features extracted from each window. Personal models, where training and testing data originate from a single user, were developed. In our study, 44 personal models were built (data from 6 participants were discarded due to missing data and recording errors), each evaluated using stratified 10-fold cross-validation repeated 10 times. For each participant, we obtained an average of 403.29 (55.62) samples labeled as happy, 403.67 (51.46) samples labeled as sad, and 402.93 (50.24) samples labeled as neutral. Among the 44 personal models, 16 were from Condition 1 (watch a movie and then walk), 14 were from Condition 2 (listen to music and then walk), and 14 were from Condition 3 (listen to music and then walk).

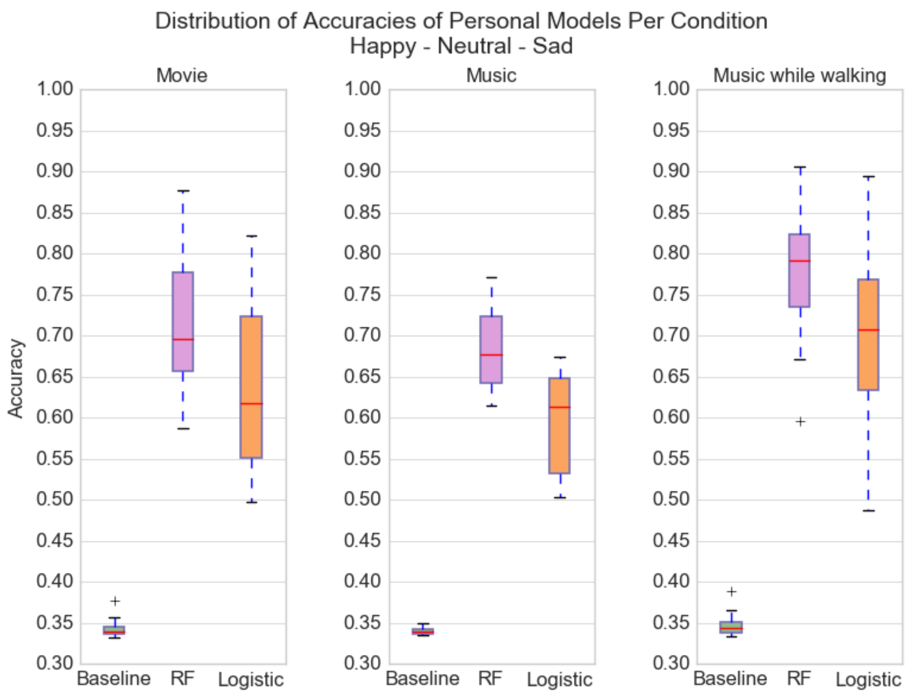
3.3 Validation

An experiment was conducted to investigate the influence of neighborhood bias on model evaluation using random cross-validation. In this experiment, a 10-fold cross-validation was performed for each personal model, with the testing fold in each iteration containing either a contiguous happy data block or a contiguous sad data block. The aim was to ascertain, with increased confidence, whether the classifiers were learning patterns associated with emotions rather than simply distinguishing between different walking periods. This validation approach considered neighborhood bias, a factor that, if overlooked, can result in overly optimistic performance estimates. The results, illustrated in Figure 5 and Table 6, indicate a decline in accuracies across all conditions compared to those obtained using random cross-validation. Although personal models outperformed personal baselines, the overall accuracy remained modest. Used Leave-One-Out Cross-Validation to see how well does the model learn patterns from other users and with what accuracy can predict on a different user.

3.4 Conclusion

Logistic regression performed poorly across all conditions when using leave one out validation method, showing that using data from different users to do emotion recognition on a different user is not possible with current features and logistic regression. Low accuracies across all conditions show that the behavior from user to user varies considerably, even when performing a similar action. Owing to the small number of users per condition (¡18), data may not be enough to make accurate predictions for users not included in the training set. However, it also highlights a limitation in their modeling approach, in that different features or more advanced models may be necessary to generalize across users. Ideally, deployment of an app should include an initial data collection and calibration phase, which can be used to build a high accuracy personal model for each user.

Classical feature-based approach to emotion recognition was used till here, but it requires domain-specific, expert knowledge about the sensors and signals to extract meaningful and informative features. Alternatively, an end-to-end concept can be used, which omits signal preprocessing and feature extraction, i.e., the acquired raw signals are directly passed to the deep learning architectures assuming they will be able to extract the essential information on their own. The next steps are to use more signals gathered from the smart watch such as: EDA, BVP, ACC, HR, IBI, TEMP to train a neural network.



Distribution of Accuracies of Personal Models Per Condition
Happy - Neutral - Sad
Movie
Music
Music while walking
Accuracy
1.00
0.95
0.90
0.85
0.80
0.75
0.70
0.65
0.60
0.55
0.50
0.45
0.40
0.35
0.30
Baseline
RF
Logistic

4 References

[1] Quiroz, Juan Carlos, Elena Geangu, and Min Hooi Yong. ”Emotion recognition using smart watch sensor data: Mixed-design study.” JMIR mental health 5.3 (2018): e10153.

[2] S. M. A. Abdullah, S. Y. A. Ameen, M. A. M. Sadeeq, and S. Zeebaree, “Multimodal Emotion Recognition using Deep Learning”, JASTT, vol. 2, no. 02, pp. 52-58, Apr. 2021.

[3] Baltrušaitis, Tadas, Chaitanya Ahuja, and Louis-Philippe Morency. ”Multimodal machine learning: A survey and taxonomy.” IEEE transactions on pattern analysis and machine intelligence 41.2 (2018): 423-443.

[4]Dzedzickis, Andrius, Artūras Kaklauskas, and Vytautas Bucinskas. ”Human emotion recognition: Review of sensors and methods.” Sensors 20.3 (2020): 592.

[5]Saganowski, Stanisław, et al. ”Emognition dataset: emotion recognition with self-reports, facial expressions, and physiology using wearables.” Scientific data 9.1 (2022): 158.

[6]Shu, Lin, et al. ”Wearable emotion recognition using heart rate data from a smart bracelet.” Sensors 20.3 (2020): 718.

[7] B. Zhao, Z. Wang, Z. Yu and B. Guo, ”EmotionSense: Emotion Recognition Based on Wearable Wristband,” 2018 IEEE SmartWorld, Ubiquitous Intelligence Computing, Advanced Trusted Computing, Scalable Computing Communications, Cloud Big Data Computing, Internet of People and Smart City Innovation

[8]Naveed Ahmed, Zaher Al Aghbari, Shini Girija, A systematic survey on multimodal emotion recognition using learning algorithms, Intelligent Systems with Applications, Volume 17, 2023, 200171, ISSN 2667-3053

[9]Cimtay, Yucel, Erhan Ekmekcioglu, and Seyma Caglar-Ozhan. ”Cross-subject multimodal emotion recognition based on hybrid fusion.” IEEE Access 8 (2020): 168865-168878.

[10]S. Saganowski, B. Perz, A. G. Polak and P. Kazienko, ”Emotion Recognition for Everyday Life Using Physiological Signals From Wearables: A Systematic Literature Review,” in IEEE Transactions on Affective Computing, vol. 14, no. 3, pp. 1876-1897, 1 July-Sept. 2023

[11]Saganowski, Stanisław. ”Bringing emotion recognition out of the lab into real life: Recent advances in sensors and machine learning.” Electronics 11.3 (2022): 496.

[12]Moontaha, Sidratul, Franziska Elisabeth Friederike Schumann, and Bert Arnrich. ”Online learning for wearable eeg-based emotion classification.” Sensors 23.5 (2023): 2387.

[13]Imani, Maryam, and Gholam Ali Montazer. ”A survey of emotion recognition methods with emphasis on E-Learning environments.” Journal of Network and Computer Applications 147 (2019): 102423.

[14]V. Dissanayake, S. Seneviratne, R. Rana, E. Wen, T. Kaluarachchi and S. Nanayakkara, ”SigRep: Toward Robust Wearable Emotion Recognition With

Contrastive Representation Learning," in IEEE Access, vol. 10, pp. 18105-18120, 2022, doi: 10.1109/ACCESS.2022.3149509.

[15]Sharma, Garima, and Abhinav Dhall. "A survey on automatic multimodal emotion recognition in the wild." Advances in data science: Methodologies and applications (2021): 35-64.