IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Date of publication 22.02.2024, date of current version 22.02.2024
Digital Object Identifier

Exploring Physiological Correlation of Affective States Using RECOLA Dataset

LUKASZ SZARECKI, JAKUB KONIUSZEWSKI, GIUSEPPE BOCCIGNONE, IEEE Senior Member
Corresponding author: Giuseppe Boccignone (e-mail: giuseppe.boccignone@unimi.it)

This research was funded by InSecTT (www.insectt.eu) project that has received funding from the ECSEL Joint Undertaking (JU) under grant agreement No 876038. The JU receives support from the European Union's Horizon 2020 research and innovation programme and Austria, Sweden, Spain, Italy, France, Portugal, Ireland, Finland, Slovenia, Poland, Netherlands, Turkey. The document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

ABSTRACT This paper explores the potential of the RECOLA dataset for investigating the physiological correlation of affective states through biosignals exploration. We propose to analyze the Biosignals module of the RECOLA dataset to uncover patterns related to valence and arousal. Specifically, we aim to examine discernible patterns in electrocardiogram (ECG) and electrodermal activity (EDA) signals that correlate with the positive or negative valence and low or high arousal of emotional states. Our analysis will focus on identifying associations between physiological signals, such as heart rate variability and skin conductance, and emotional experiences. By investigating these patterns, our aim is to enhance our understanding of the physiological correlates of affective states using physiological data.

INDEX TERMS Physiological signals, Affective states, Biosignals exploration, Valence, Arousal, Electrocardiogram (ECG), Electrodermal activity (EDA), Machine learning, RECOLA dataset

I. INTRODUCTION

In recent years, there has been a growing interest in understanding the intricate relationship between physiological signals and affective states. Affective states, encompassing a spectrum of emotions and moods, are integral components of human experience and play pivotal roles in interpersonal interactions, decision-making, and overall well-being. Traditional methods of emotion recognition have often relied on observable cues such as facial expressions, tone of voice, and body language. However, emerging research suggests that physiological signals, including electrocardiogram (ECG) and electrodermal activity (EDA), offer valuable insights into the underlying physiological correlates of affective states.

Against this backdrop, we turn our attention to the RECOLA dataset, a rich repository of physiological recordings capturing biosignals associated with affective states. This dataset provides a unique opportunity to delve deeper into the physiological manifestations of emotional experiences, particularly focusing on patterns related to valence and arousal, which are fundamental dimensions of affective states.

Motivated by the potential of biosignals exploration, our proposed project aims to analyze the physiological recordings from the Biosignals module of the RECOLA dataset. Specifically, we seek to identify discernible patterns in ECG and EDA signals that may correlate with the positive or negative valence of emotional states. By leveraging machine learning techniques and advanced signal processing algorithms, we aim to uncover subtle associations between physiological signals and emotional experiences.

Our investigation holds promise for shedding light on the complex interplay between physiological responses and affective states. For instance, certain patterns in heart rate variability or skin conductance may serve as indicators of positive or negative emotional experiences. Understanding these patterns could pave the way for the development of more accurate and robust emotion recognition systems, with implications for various domains including psychology, healthcare, and human-computer interaction.

In our paper, we take a detailed approach to studying emotional states, using the continuous valence and arousal labels provided by the RECOLA dataset. Rather than focusing solely on distinct emotion categories, we delve into the spectrum of emotional experiences. By leveraging the extensive data in the RECOLA dataset, our aim is to enhance our understanding of the physiological connections to emotional states. This endeavor contributes to the ongoing development of emotion research and technology

Overall, our project represents a significant step towards unraveling the intricate relationship between physiological signals and affective states, with potential implications for enhancing our understanding of human emotions and informing the development of innovative applications in emotion recognition and affective computing.

II. PHYSIOLOGICAL SIGNALS & AFFECTIVE STATES

A. EDA, ECG

Electrodermal Activity (EDA) and Electrocardiogram (ECG) are two crucial physiological signals often studied in the context of emotional states and arousal levels.

EDA, also known as Galvanic Skin Response (GSR), measures the electrical conductance of the skin, which is influenced by sweat gland activity. It serves as a reliable indicator of sympathetic nervous system arousal and emotional responses. Changes in EDA reflect shifts in autonomic arousal, with higher levels typically associated with increased emotional arousal and stress.

On the contrary, ECG captures the heart's electrical activity throughout time, offering critical insights into cardiac performance and unveiling patterns related to emotional arousal and stress. Specifically, the RR interval, denoting the duration between successive heartbeats, is a pivotal metric analyzed for evaluating autonomic nervous system activity and emotional regulation. Elevated HRV, indicative of increased variability in these time intervals, typically correlates with superior emotional regulation and the ability to adapt to stressors.

By analyzing EDA and ECG signals from the RECOLA dataset, we can explore how these physiological measures correlate with the valence and arousal dimensions of emotional states. Specifically, we can investigate whether certain patterns in EDA, such as rapid fluctuations in skin conductance, or in ECG, such as changes in heart rate variability, are associated with positive or negative valence and varying levels of arousal. Identifying such patterns can enhance our understanding of the physiological underpinnings of emotional experiences and potentially contribute to the development of effective emotion recognition systems and interventions for emotional well-being.

B. VALENCE, AROUSAL

Valence and arousal are two key dimensions used to describe emotional experiences and are fundamental concepts in affective psychology. Understanding these dimensions can provide valuable insights into how emotions are experienced and expressed.

Valence: Valence refers to the pleasantness or unpleasantness of an emotional experience. It represents the positive or negative quality of an emotion. Emotions with positive valence are those that evoke feelings of pleasure, happiness, or satisfaction, while emotions with negative valence are associated with feelings of displeasure, sadness, or fear. For example, joy and love are emotions typically characterized by positive valence, while sadness and anger are emotions characterized by negative valence.

Arousal: Arousal refers to the intensity or activation level of an emotional experience. It reflects the physiological and psychological activation associated with an emotion. Emotions with high arousal are intense and stimulating, often accompanied by increased heart rate, rapid breathing, and heightened alertness. Examples of high-arousal emotions include excitement, anger, and fear. Conversely, emotions with low arousal are calmer and more subdued, characterized by decreased physiological activation. Examples of low-arousal emotions include relaxation, contentment, and sadness.

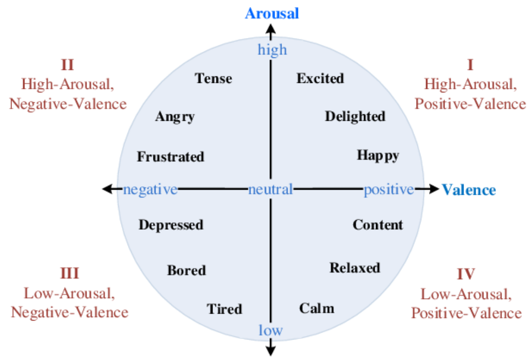

FIGURE 1. Two dimensional valence/arousal space

Understanding the interplay between valence and arousal can provide valuable insights into emotional states and their physiological correlates. By examining patterns in physiological signals such as ECG and EDA, which are known to be influenced by emotional arousal and valence, we can explore how these signals correspond to different emotional experiences.

III. SYSTEM MODEL

Valence/Arousal recognition models in Fig. 2 are constructed as binary classification models. Each model is customized based on input data, undergoing feature extraction before classification. The dataset is split into training and testing subsets, with training data exclusively instructing models. This study evaluates five binary classification algorithms: decision tree (DT), random forests (RF), K-nearest neighbor (KNN), gradient boosting (GBA), and adaptive boosting (ABA). Effectiveness of these models is then measured using designated testing data. We will consider models tailored to classify levels of arousal and valence, distinguishing between low and high states. Specifically, for valence models, one will exclusively utilize ECG features, another solely EDA features, while the third model will incorporate both feature types. Similarly, separate models will be developed for arousal classification.

A. DATASET

In the conducted research, biosignals data were used from the RECOLA dataset exclusively [4], which includes two parts: physiological recordings and features. The physiological recordings consist of Electrocardiogram (ECG) and electrodermal activity (EDA), collected using a Biopac MP150 unit and stored in a CSV file. The dataset also contains

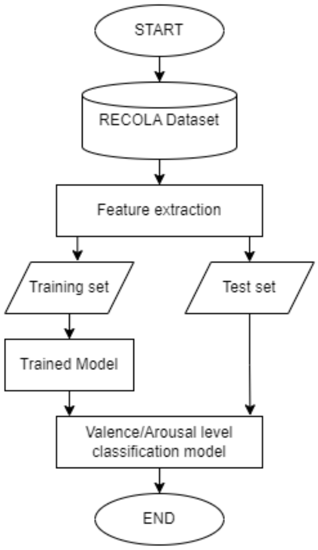

FIGURE 2. **The valence/arousal level classification process**

annotations for socio-affective behaviors performed by six assistants, three males, and three females, using the ANNEMO web-based annotation tool. These annotations provide valuable insights into affective behaviors, specifically arousal and valence, with a frame rate of 40ms. This comprehensive use of biosignals data from RECOLA allowed the authors to study the details of socio-affective behaviors in their research.

B. SIGNAL PROCESSING

In the extensive analysis of the RECOLA dataset, both valence and arousal values from examinations were thoroughly examined and visualized through charts. The dataset, comprised of annotations data, features single records lasting 300 seconds, each containing 7500 samples of valence and arousal. Assessments were contributed by six judges (three females and three males) for each of the 18 examinees. In terms of valence, the examination-wide mean value was determined to be 0.094, indicating a distinct tendency among judges to assign more positive valence values. This observation provides valuable insights into the socio-affective behaviors recorded during the examinations. In contrast, the analysis of arousal values revealed a different pattern. The total mean arousal of all examinees during the entire examination was found to be -0.007, suggesting no clear observed tendency among judges. This contrast underscores the diversity in the assessment of arousal values across the examined dataset. Figures 3. and 4. highlight the considerable variability in valence and arousal data among different examinees. To capture individual differences, we set specific thresholds for each examinee, defining what counts as high or low values.

Figures 5. and 6. showcase the results of this threshold-based classification, offering a clear distinction between high and low values for arousal and valence.

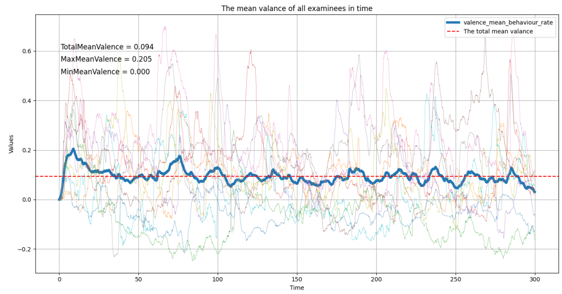

FIGURE 3. **The mean valence for all files in dataset**

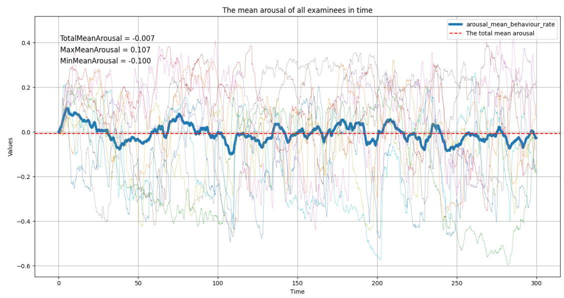

FIGURE 4. **The mean arousal for all files in dataset**

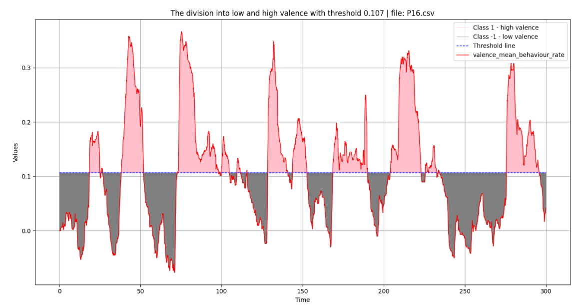

FIGURE 5. **The division into low and high valence based on mean_behaviour_rate of valence for single file**

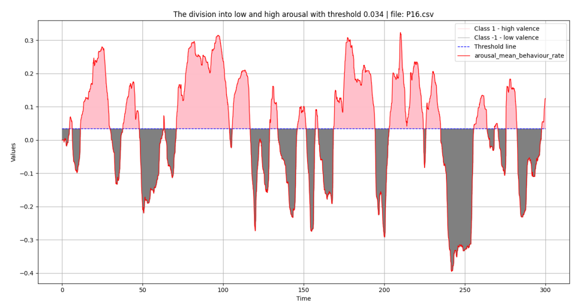

FIGURE 6. **The division into low and high arousal based on mean_behaviour_rate of arousal for single file**

C. FEATURE EXTRACTION

EDA

After analyzing the EDA signal, focus is on detecting noticeable peaks within the signal. This approach is motivated by the understanding that changes in skin conductivity, reflected in the peaks of the EDA signal, are often associated with shifts in emotional states. By identifying these peaks, it is aimed to capture significant variations in the electrodermal response that may indicate changes in arousal and valence levels.

Capturing these peaks allows for the quantification of the intensity and timing of physiological responses to emotional stimuli, providing valuable information for understanding the dynamics of emotional experiences. By incorporating this feature into the learning model, we can enhance its ability to recognize and classify different emotional states based on patterns of electrodermal activity. This feature extraction method enables the model to leverage the relationship between changes in skin conductivity and arousal/valence, improving its overall accuracy and robustness in level classification.

ECG

After analyzing the ECG signal, the focus was on extracting a key feature known as the RR interval. The RR interval represents the time between successive R peaks in the ECG waveform, reflecting the duration of one complete cardiac cycle. This feature was chosen because changes in the RR interval are closely linked to alterations in autonomic nervous system activity, which can be indicative of shifts in emotional arousal and valence levels. By tracking variations in the RR interval, patterns related to sympathetic and parasympathetic modulation of the heart rate can be discerned, providing insights into the individual's emotional state

IV. CLASSIFICATION

A. DECISION TREES (DT)

Authors created the first classification model using a decision tree with default parameters based on the Classification and Regression Trees (CART) [1] algorithm from the scikit-learn library. This algorithm can be employed for both classification and regression and involves constructing a tree structure by recursively partitioning the data based on the Gini impurity as the criterion for splitting nodes. The model is trained with each root node having a condition created using one of the explanatory variables, according to which the data is separated.

During the experiments, we investigated the impact of varying maximum tree depths. Through a five-fold cross-validation process, the highest accuracy was observed for depths within the range (19,20,21).

B. RANDOM FOREST (RF)

A natural extension of decision trees is random forests. They allow for a natural way to minimize the compromise between bias and variance, as random forests are inherently generated by creating multiple trees using the Bootstrap method. When constructing trees, one can choose to have them built from a subset of all classes, introducing additional randomness. Consequently, each model within the forest differs from others because each is created from a different subset of data determined using the mentioned methods. This implies that during the creation process of each tree, it assigns different weights to the data features. The strength of these models lies in the ensemble voting method, facilitating unanimous decision-making.Random forests are an extension of the decision tree method, aimed at addressing the main drawback of trees, which is overfitting to the training portion of the dataset. [2]

In the case of decision trees and their maximum depth, increasing the depth results in a higher number of branches, consequently requiring more time for its creation. Another parameter for restriction is the number of trees in the forest – the more trees, the computationally more complex it becomes, but also more diverse trees can contribute to the decision-making process, reducing variance. However, beyond a certain threshold, the advantages of having more trees become negligible. It was crucial to determine the optimal number of trees that achieves the highest accuracy while simultaneously limiting computational power."

In the experimental findings, it emerged that optimal results were obtained when utilizing a random forest with 150 trees and setting the maximum depth of each tree to 22. To utilize this model, the scikit-learn Python library was imported.

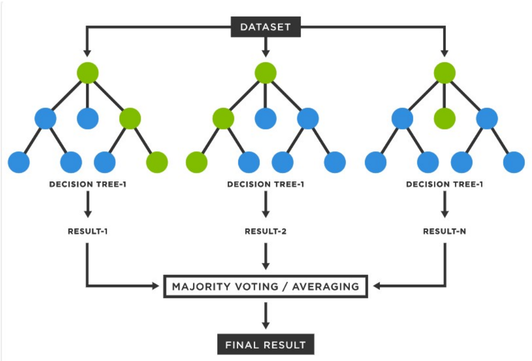

FIGURE 7. Example of Random Forest structure [3]

C. K-NEAREST NEIGHBORS ALGORITHM (KNN)

The next classification model was constructed using the k-Nearest Neighbors (KNN) algorithm with default parameters. Employed exclusively for classification tasks, KNN from the scikit-learn library relies on the principle of assigning labels to instances based on the majority class among their k-nearest neighbors. This algorithm does not involve explicit training; instead, it memorizes the entire dataset for immediate predictions during inference. The proximity between instances is measured using a distance metric, typically Euclidean distance, influencing the decision boundaries of the model.

During the experiments, we found that the best results were obtained with the default configuration for KNN.

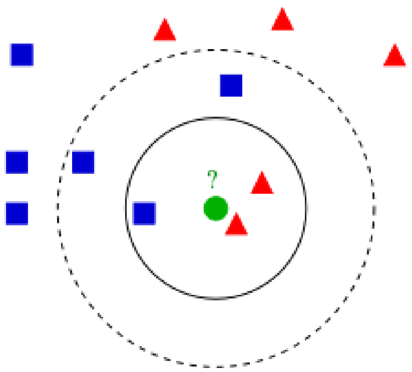

FIGURE 8. **Example of k-NN classification [5]**

D. GRADIENT BOOSTING ALGORITHM (GBC)

The following classification model was constructed using the Gradient Boosting algorithm with default parameters, specifically designed for classification tasks. Accessible in the scikit-learn library, Gradient Boosting iteratively boosts the performance of weak learners, commonly decision trees, refining accuracy. This algorithm, known for its iterative learning approach, corrects errors from previous weak learners. Known for understanding complex data connections, Gradient Boosting is a strong solution for classification tasks, making it a good choice for managing complex datasets and enhancing model accuracy.

Throughout the experiments, we observed that optimal outcomes were achieved when using 400 estimators. Further adding this attribute didn't contribute to model's increased accuracy.

E. ADAPTIVE BOOSTING ALGORITHM (ADA)

Adaptive Boosting Algorithm (AdaBoost) is an ensemble learning method primarily utilized for classification tasks. It iteratively constructs a sequence of weak learners, typically shallow decision trees or stumps, by focusing on instances that were previously misclassified. AdaBoost assigns weights to each instance, with misclassified instances receiving higher weights, thus influencing subsequent iterations. It combines the predictions of weak learners using a weighted majority vote to form a strong learner. This iterative process results in a accurate classification model capable of handling complex datasets and improving performance over time.

During experiments with AdaBoost that was accessible through the scikit-learn library, it was observed that the optimal number of estimators(stumps) was approximately 500. Beyond this point, increasing the number of estimators did not significantly enhance the model's accuracy.

V. RESULTS AND CONCLUSION

Following figures display the classification accuracy using five distinct machine learning models for RECOLA dataset, respectively.

Precision has been used as a classification metric, as it is known - it is a measure of the accuracy of positive predictions made by the model and is calculated as the ratio of true positives (correctly predicted positive instances) to the sum of true positives and false positives (instances wrongly classified as positive). This way, whilst examining arousal/valence levels, low and high scales of these levels are being classified.

	ADA	DT	GBC	KNN	RF
EDA	0.65	0.69	0.67	0.68	0.69
ECG	0.58	0.58	0.58	0.57	0.58
Both	0.65	0.74	0.7	***0.76***	0.74

FIGURE 9. **Prediction accuracy of applying different models to RECOLA - classyfying high arousal levels by EDA,ECG and both signals**

	ADA	DT	GBC	KNN	RF
EDA	0.64	0.62	0.65	0.62	0.65
ECG	0.54	0.5	0.55	0.48	0.52
Both	0.65	0.71	0.71	0.71	***0.76***

FIGURE 10. **Prediction accuracy of applying different models to RECOLA - classyfying low arousal levels by EDA,ECG and both signals**

	ADA	DT	GBC	KNN	RF
EDA	0.64	0.64	0.64	0.62	0.68
ECG	0.54	0.5	0.54	0.48	0.56
Both	0.64	0.71	0.68	0.72	***0.76***

FIGURE 11. **Prediction accuracy of applying different models to RECOLA - classyfying high valence levels by EDA,ECG and both signals**

	ADA	DT	GBC	KNN	RF
EDA	0.66	0.66	0.66	0.66	0.7
ECG	0.58	0.56	0.58	0.56	0.52
Both	0.66	***0.75***	0.7	***0.75***	0.73

FIGURE 12. **Prediction accuracy of applying different models to RECOLA - classyfying low valence levels by EDA,ECG and both signals**

Results demonstrated varying accuracy levels across different models - EDA signals showed more promising results compared to the ECG. The Random Forest(RF) utilizing the EDA signal tends to attain best results, representing the most accurate prediction outcomes up to 70% in this investigation. When we take under consideration the model in which features from both signals are combined into a single feature set, we obtain more promising accuracy results up to 76%. In this way, the model utilizes more diverse information from both EDA and ECG features, making it the superior one. After testing many approaches with various classification models it became evident that the most promising performers in this case were Decision Tree, K-Nearest Neighbors (KNN), and Random Forest. We opted not to include RNN, linear regression, and Naive Bayes in our tables and analyses due to their subpar performance and weak results during testing. Consequently, these models were not considered for further evaluation.

To understand gotten differences, one could say that EDA signals directly reflect the activity of the sympathetic nervous system, which is intricately involved in the body's physiological response to arousal and valence levels. In contrast, while ECG signals also provide insights into physiological functioning, they may not capture arousal and valence levels as directly or sensitively as EDA signals. As individuals experience emotions such as excitement, anxiety, or stress, their sympathetic nervous system activity increases, leading to changes in sweat gland activity and subsequently in EDA signals. These changes can occur rapidly and are often detectable even for subtle emotional shifts.

Expanding the scope of the study to include additional datasets could yield more comprehensive insights and potentially enhance the accuracy of the model predictions. Relying solely on the RECOLA dataset, which comprised only 18 CSV files, may not provide a complete representation of the diverse range of socio-affective behaviors and emotional responses. Furthermore, the variability in the levels of arousal and valence rated by different individuals within the dataset suggests potential inconsistencies and limitations in the data. Therefore, exploring additional datasets with a larger and more diverse sample size could offer a broader perspective and help validate the robustness of the model across different contexts and populations. This approach would not only improve the reliability of the model but also contribute to a more thorough understanding of affective behaviors and emotional dynamics.

Moreover, feature extraction process was simplified by utilizing a specific feature that aligned with a particular level of arousal or valence. Incorporating more sophisticated or diverse features, or utilizing advanced feature extraction techniques, could have provided a richer representation of the underlying patterns and relationships within the data. By extending the feature extraction process to include a broader range of features or more complex feature representations, the models' predictive capabilities could have been further optimized. However, it's worth noting that the authors have opted for a simplified approach to streamline the analysis process.

Therefore, exploring additional datasets with a larger and more diverse sample size could offer a broader perspective and help validate the robustness of the model across different contexts and populations. This approach would not only improve the reliability of the model but also contribute to a more thorough understanding of affective behaviors.

REFERENCES

[1] Wei-Yin Loh. Classification and regression trees. Wiley interdisciplinary reviews: data mining and knowledge discovery, 1(1):14–23, 2011.
[2] Leo Breiman. Random forests. Machine learning, 45(1):5–32, 2001.
[3] https://www.spotfire.com/glossary/what-is-a-random-forest. (15.02.2024).
[4] RSS13: F. Ringeval, A. Sonderegger, J. Sauer and D. Lalanne, Introducing the RECOLA Multimodal Corpus of Remote Collaborative and Affective Interactions, 2nd International Workshop on Emotion Representation, Analysis and Synthesis in Continuous Time and Space (EmoSPACE), in proc. of IEEE Face & Gestures 2013, Shanghai (China), April 22-26 2013.
[5] https://en.wikipedia.org/wiki/File:KnnClassification.svg. (15.02.2024).