



Emotion Classification with Wearables

Name

Institutional Affiliation

Emotion Classification with Wearables

I. Introduction

The importance of emotions and related psychological processes is immeasurable in terms of the effect it has on the daily lives of people. The topic of emotions used to be the prerogative of many spheres of human activity like philosophy or religion. Only during the last two centuries, with the emergence of psychology, this notion has become the subject of scientific research. Even when the changing character of psychology as a scientific discipline is taken into account, it can be said that researchers in this area of studies have made the most notable progress in understanding human emotions. With the rapid development of technologies, it became possible to bring objective and accurate measurement methods for registration of the events that used to be studied exclusively using the phenomenological approach. While some simple ways of detecting the change in the emotional state have in use for more than a hundred years, only the recent achievements in computer science, machine learning, and engineering made it possible for researchers to look deeper into the mechanisms of emotional processes.

While the topic became a subject of significant scientific interest, the area of knowledge on how emotional processing occurs comprises a number of valuable practical applications of the information mentioned. Currently, the popular demand for being aware of one's emotions stems from psychological counseling, self-help and personal growth industry, supportive medicine, and many other fields. The main issue related to fulfilling the demand is connected to limitations that exist in the domain of technical provision of the emotion registration process. Emotions, just like any other phenomena of psychology, are generated by the brain. Researchers use expensive equipment to study the neural processes. However, such devices are few and mostly available to either large research facilities or clinics. In this way, the portable, wearable, and low-cost products are supposed to be the solution.

However, brain activity is extremely complex, and it is hard to register the needed information using the limited capacities of the devices available. Those obstacles are imposed by the necessities of equipment to have good usability and user-friendliness. In order to address the issues outlined above, there was an attempt to create the neuropsychological and mathematical framework for the future device that would combine the accuracy of the medical-tier equipment with the availability and accessibility of the smartphone. To achieve this goal, the given work will present the new

method for emotion classification that has the potential of being incorporated into affordable wearable devices for everyday use. This includes the assessment of the parameters related to heart functioning combined with the methods of machine learning. In this way, it would be possible to create a simple device that can tell valuable information on one's emotional condition. While not having a need for conducting a complicated process of brain scanning or similar procedures, the wearable with a function of tracking and analyzing the information on heartbeat can still serve as a valid tool for an indication of emotion. Naturally, it would not be possible to make a distinction between the complex emotions like grief or spleen, its valence and relevance can be shown with a significant measure of accuracy.

II. Related Work

Before beginning any work on the emotion registration, recognition, or classification, it is necessary to have proper methodological guidance in the subject. As many other subfields of psychological science, the psychology of affective states is multifaceted and represented by the numerous theoretical approaches to the explanation of the subject. There is still a debate in the professional literature regarding the question of whether emotions are universal in all cultures or not. The opponents of the universalist position argue for the numerous differences in the cultural norms of expressing oneself. Also, they often present an argument that the specific name behind a certain perceived emotional state may be unique for a given cultural context and not have any word analogs in other languages. However, as the present work is explicitly aimed at creating a product that can register emotions, it is necessary to have a theoretical basis that at least recognizes a certain measure of similarity in the way human beings run the related processes. For these purposes, it was decided to adhere to theoretical models that focus on finding the common neuropsychological and physiological basis for the emotion-related events regardless of their name as the classification is based not on the linguistic differentiation, but the observed objective processes in a human body.

It is still difficult to say exactly what an emotional reaction is; some scientists argue that this is only a product of physiological processes, while others state that the latter is a result of emotions. However, in this study, the cause-and-effect relationship is of a lesser significance. It is important to mention that a person is not able to hide the emotional reaction. Indeed, one has an option to suppress one's actual feelings, but it is almost impossible to hide the emotions completely. The polygraph is one of the means used when it is necessary to find out the truth, and it is believed that it is difficult to cheat on the polygraph test, at least for an unprepared person. However, the polygraph function of defining the truth and lies is rather relative. The truth is what a person believes to be so, and one can believe in anything. A lie is an attempt to hide what one remembers and knows. In any way, it remains possible to deceive the polygraph, and its use is strongly limited by laboratory conditions. It means that it is necessary to develop systems capable of determining the relation of an object to something, regardless of the conditions of the test, of whether the subject knows about the test being conducted or not, and of any external influences.

Some of the latest studies in the area suggest that there is a connection between the experiencing of emotions and the functional state of the autonomous nervous system. The changes in the activity of the latter can be measured in many different ways. For example, the heart rate variability (HRV) was shown to have a consistent correlation with the emotional state of the respondents. The research in the area indicates that the resting-state rate of heartbeat is not regular as was previously believed. The beat-to-beat changes in heart rate occur through interaction between the sympathetic and parasympathetic components of the autonomous nervous system. The parasympathetic system slows the heart down, and the sympathetic component acts in the opposite direction. The changes in heart rate variability allow to understand whether a person experiences an emotion of a certain magnitude or is being calm even if the respondent deliberately

tries to hide one's inner feelings.

Researchers are turning to the use of an electroencephalogram (EEG), which is devoid of the above disadvantages. With regards to the dynamical proceeding of emotional stimuli, EEG has an advantage even over such modern methods like functional magnetic resonance imaging (fMRI) and positron emission tomography (PET), due to sufficient temporal resolution. In the works devoted to the study of emotions and the allocation of indicators of emotions experienced in EEG, two main methodological approaches are used. The first approach is based on the analysis of changes in the components of evoked potentials upon presentation of stimuli. The stimulation in such studies is usually produced by the effects of different modalities. The results of the studies were ambiguous and were primarily determined not by a factor in the development of an emotional reaction, but by the conditions of the experiments. Another approach to the study is the analysis of the spectral and coherent characteristics of the EEG when the subjects perform various tasks related to emotional experiences. Despite the relatively large amount of accumulated factual material on changes in the EEG of a person that experiences intense emotions, the data obtained are often challenging to interpret, primarily because of the need to distinguish these changes from similar ones that occur during non-emotional tasks, especially if the matter is not simple. For instance, it can be regarded in the case of presenting rather complex emotiogenic tasks that include cognitive components. The latter sometimes play the most significant role. Thus, the need for further search is dictated, on the one hand, by the undoubted practical importance of identifying objective indicators of the sign and intensity of emotions, and, on the other, by the uncertainty and inconsistency of the data obtained in this area. As emphasized by many authors, both physiological and psychological, emotional manifestations are complex and sophisticated. Lamerun notes that emotion is a complex phenomenon that includes at least three aspects:

1. Experienced or even perceived feeling (subjective phenomenology of emotions);
2. Processes occurring in the endocrine and autonomic nervous system of the body (visceral phenomenology of emotions), as well as the expression of feelings, intonation, gestures, and postures (behavioral phenomenology of emotions);
3. Changes in the central nervous system and brain activity (central phenomenology of emotions).

The topic of emotion recognition using various physiological parameters is not new, as it is seen from the works summed up above. It can be noted that there is a perspective direction for the new research. The multimodal dataset has already been studied thoroughly. However, the results of such research are not very useful for designing an affordable wearable device that only has a heartbeat sensor. Affective states can be classified using the characteristics of valence and arousal focuses. These two metrics show the degree to which the emotional processes are present in a person's affective experience. The valence stands for determining whether an emotion is pleasant or unpleasant, whereas the arousal signifies the measure of activation or deactivation. In engineering-psychological and biomedical research, especially when assessing the level of the patient's psycho-emotional tension, human electrophysiological indicators are widely used. At the moment, there are many methods to identify and fix the electrophysiological indicators of a person. The correct choice of methods, the adequate use of its indicators are the conditions necessary for conducting a successful psychophysiological study. This area of research is significant both in medicine and engineering.

Today, among the most diagnostically significant electrophysiological methods, the following can be singled out: electroencephalogram study, skin galvanic reaction, temperature, electrooculogram, electromyogram, plethysmogram. Also, electrocardiography (ECG) is often used. These methods allow recording such parameters as muscle arousal, rapid heartbeat, blood outflow from the skin surface of a person, brain activity, etc. According to research conducted by psychological services,

these methods also allow recording the change in the emotional state of a person. Spectral methods for the analysis of HRV are now very widespread. The examination of the power spectral density of oscillations gives information about the distribution of power depending on the frequency of oscillations. Moreover, HRV is a highly sensitive indicator of the dynamic processes of realization of psycho-emotional states in a person, arising as a result of radical quantitative and qualitative situational reorganizations of the personality structure of the subject in the processes of mental adaptation. It is believed that the fluctuations (variations, deviations) of the heart rhythm, manifested in the “scatter functions” provide the necessary information to assess the quality of mental adaptation processes. For example, the involvement of the limbico-reticular complex and the frontal-temporal structures of the right hemisphere in these processes was revealed by the indicators of the dynamics of slow waves of the first order in the spectrogram of the heart rhythm. Informativeness of HRV indicators was found in the diagnosis of the degree of anxiety and depressive disorders, which are formed as a result of stagnant reverberation excitations at the level of limbic-reticular brain structures.

III. Proposed Methods

First, IBI time series was extracted from the ECG signal using a combined adaptive threshold. Only the last 60 seconds of ECG were used due to the variable duration of the stimuli clips and to make sure only one emotion is presented in the data. This provides a sparse vector of the time difference between heartbeats. As mentioned before, IBI is already a well-validated measurement for affective states classification. Next, the IBI vectors were normalized to boost convergence of the model. Dividing by 2000 is enough to make sure the IBIs are in the range of $[0,1]$. Vectors were also flipped to make sure the last beats are always in a fixed position, assuming they correlate better with the affective state. Lastly, IBI vectors were zero padded to the maximum length of the longest IBI vector (101 beats). A flow diagram of the overall pre-processing is shown in Figure 1.

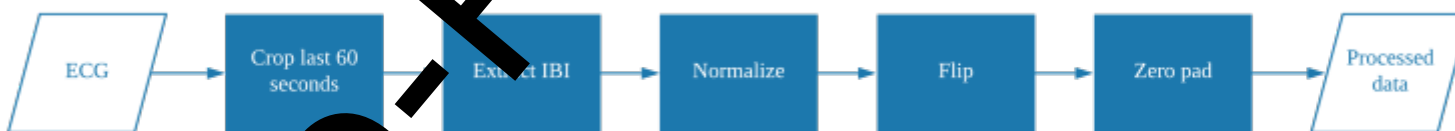


Figure 1. Flow Diagram.

A new architecture pattern “convolutional block” is introduced as a set of connected 1D convolutional layers. Each layer is followed by batch normalization and an activation function. The feature extraction network is built by connecting multiple convolutional blocks with pooling and dropout layers between them. After extracting the features, a final pooling layer is connected to reduce and flatten the features vector before feeding it to the classification network. This network is built by fully connected layers where there is a batch normalization, an activation function, and a dropout layers between each of them. The last layer of the network is a single unit with a sigmoid activation to output a binary classification of the affective state. Two independent models were trained, one for valence and the other for arousal classification. Together the two models can classify the affective state. The proposed architecture for affective states classification was implemented with Python and the Keras framework.

For training the models, a dataset was created. The ECG input was processed based on the method discussed in Section II-A. The labels were extracted from the SAM of the relevant stimuli clip. If the value (valence/arousal) is greater than 3, a label of 1 or 0 was set. Film clips with

the value 3 were ignored as they indicate a neutral state; furthermore, data were truncated to make sure the dataset is balanced and there is no bias. A parameter space was defined to find the hyperparameters which will yield the best performance. The parameter space includes the number of convolutional blocks, number of convolutional layers for each block, kernel size, filter size, dropout rate, activation function, pooling layer, whether to use batch normalization, final pooling layer, number of units in the fully connected layer (if any), learning rate, and optimizer. Performance is defined as the median accuracy of 5-fold cross-validation. During the hyperparameters search, this metric was optimized. The same parameter space was used both the valence and arousal models, but two different parameter sets have eventually been chosen. The best parameters were used to train the final models. A binary cross-entropy loss function was chosen, and models were trained for 200 epochs.

IV. Results

A 5-fold cross-validation technique was used to assess the model performance, as discussed in Section II-C. The split was done based on participants compared to DREAMER authors who separated the dataset by clips. Splitting by participants is suited towards real-life applications, as most of the time, it will not possess any prior knowledge about the person but still would be able to provide an accurate prediction. Most folds consist of 5 participants, but some may only have 4 as the total number of participants is 23. Accuracy was calculated traditionally, as the ratio of correct classifications over the whole samples. Both training and validation sets are balanced; the accuracy metric does not introduce a bias towards any side. Finally, the accuracy median of the 5-fold cross-validation was used as a performance metric. We compare the performance results of the proposed method with the performance obtained initially by Stamos et al. of the dataset DREAMER. It is important to emphasize that is not a fair comparison as Stamos et al. used 10-fold cross-validation and split the set by the stimuli clips. Furthermore, they used imbalanced dataset, which might introduce bias in the accuracy metric.