Healthcare Transformation with Informatics and Artificial Intelligence
J. Mantas et al. (Eds.)

doi:10.3233/SHTI230418

Deep Learning Framework for Categorical Emotional States Assessment Using Electrodermal Activity Signals

Praveen Kumar GOVARTHAN[a,1], Sriram Kumar P [a], Nagarajan GANAPATHY [b] and Jac Fredo AGASTINOSE RONICKOM [a]

[a] *School of Biomedical Engineering, Indian Institute of Technology (BHU), Varanasi, Uttar Pradesh, India*

[b] *Department of Biomedical Engineering, Indian Institute of Technology Hyderabad, Telangana, India*

ORCiD ID: Praveen Kumar Govarthan https://orcid.org/0000-0002-6982-7327

Abstract. In this study, we attempted to classify categorical emotional states using Electrodermal Activity (EDA) signals and a configurable Convolutional Neural Network (cCNN). The EDA signals from the publicly available, Continuously Annotated Signals of Emotion dataset were down-sampled and decomposed into phasic components using the cvxEDA algorithm. The phasic component of EDA was subjected to Short-Time Fourier Transform-based time-frequency representation to obtain spectrograms. These spectrograms were input to the proposed cCNN to automatically learn the prominent features and discriminate varied emotions such as amusing, boring, relaxing, and scary. Nested k-Fold cross-validation was used to evaluate the robustness of the model. The results indicated that the proposed pipeline could discriminate the considered emotional states with a high average classification accuracy, recall, specificity, precision, and F-measure scores of 80.20%, 60.41%, 86.8%, 60.05%, and 58.61%, respectively. Thus, the proposed pipeline could be valuable in examining diverse emotional states in normal and clinical conditions.

Keywords. Convolutional neural network, electrodermal activity, emotion detection, time-frequency representations

1. Introduction

Accurate classification of human emotions is a promising tool to determine and characterize various neurological disorders associated with emotional states, such as mood, depression, anxiety, bipolar, and autism spectrum disorders [1]. According to statistics published in 2019 by WHO, 1 in every 8 people, or 970 million people worldwide, have anxiety and depressive disorders [2]. However, the prevalence rate has

[1] Corresponding Author: P.K. Govarthan, praveenkumarg.bme21@itbhu.ac.in.

increased dramatically post-COVID-19 pandemic. It is, therefore, critical to develop a potential solution for people with mental health disorders that can detect emotions automatically. Although multiple physiological modalities are explored to predict the human emotion, we have used Electrodermal Activity (EDA) signals as they cannot be altered intentionally and can provide simple, effective, low-cost, non-invasive, and continuous recordings. Further, it provides the possibility to be incorporated with consumer-grade wearables easily [1].

Studies have proved that the Time-frequency (TF) analysis provides comprehensive information on non-stationary signals like EDA. It can better describe rapid changes in signals which are not represented in the time or frequency domains separately. Due to the non-stationary and multi-component nature of EDA, an optimal TF analysis can better reveal these structures. The standard feature engineering-based classification methods involve selecting, manipulating, and transforming raw data into features that require high manual interventions and are time-consuming. Deep learning (DL) architectures can overcome the shortcomings mentioned above, automatically learn the most suitable features from the training data and yield higher classification accuracy. In this study, the Short Time Fourier Transform (STFT) based TF method and configurable Convolutional Neural Network (cCNN) based DL architecture has been used to recognize amusing, boring, relaxing, and scary emotional states [3]. The proposed approach also has the potential to bridge communication gaps for patients with non-verbal communication difficulties, thereby improving their quality of care and quality of life.

2. Methods

The proposed pipeline for the classification of emotional states is shown in Figure 1.

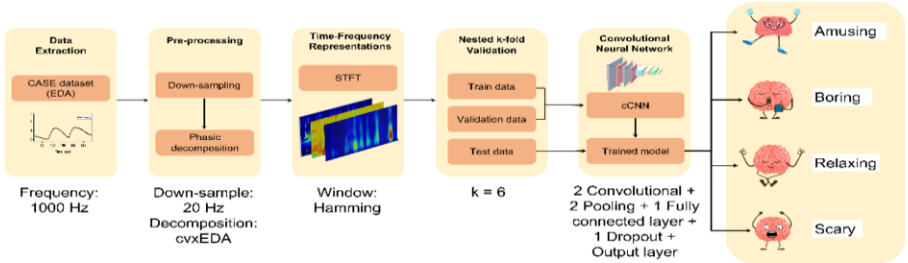
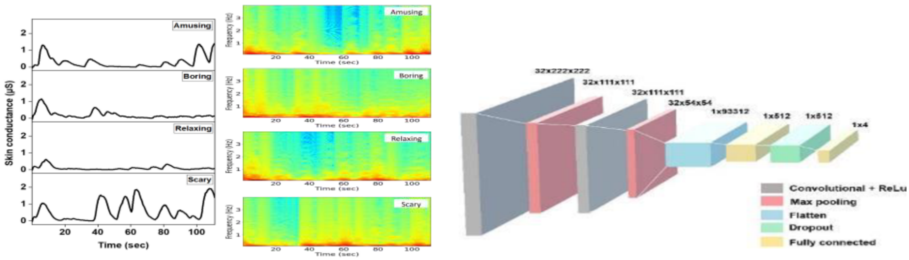


Figure 1. Flowchart of the proposed framework

The EDA considered in this study were obtained from the Continuously Annotated Signals of Emotion (CASE) dataset [4]. These signals were acquired from 30 subjects (aged: 22-37 years, sampling rate: 1000 Hz) while watching four emotional audio-video stimuli (amusing, boring, relaxing, and scary) and simultaneously reported their emotional experience using the joystick-based emotion reporting interface. EDA was down-sampled to 20Hz to conserve memory and processing time of the data without a significant risk of losing essential aspects of the signals. The down-sampled EDA were de- composed into tonic and phasic components using the convex-optimization-based EDA model (cvxEDA) algorithm, a decomposition method based on maximum a posteriori probability, convex optimization, and sparsity [5]. Further, the phasic components of EDA were subjected to STFT-based Time-Frequency Representation (TFR) [6]. Each of the (Hamming) windowed segments of the EDA signal was processed with a 2048-point fast Fourier transform. In this study, we employed the Nested 6-fold cross-validation technique to assess the model and obtain a reliable output [7]. Two convolutional, two pooling, a fully connected layer, a dropout, and an output layer were components of the proposed cCNN architecture to categorize the spectrogram into four distinct emotion classes [3].

3. Results and Discussions



(a) (b) (c)

Figure 2. (a) Representative signals, (b) Corresponding STFT under different emotional states, and (c) Proposed cCNN architecture.

Figure 2(a) displays the representative phasic components of the EDA signal of a participant under different emotional states determined by the cvxEDA algorithm. The phasic component of amusing and scary stimuli has a greater amplitude and variability, which may be explained by increased sweat gland activity during amusement and fear emotions. A similar trend can also be visualized in the respective TF images shown in Figure 2(b). These TF images are fed to the cCNN, illustrated in Figure 2(c). The results shown in Figure 3 indicate that the proposed pipeline can discriminate different emotional states with a high average classification accuracy, recall score, specificity, precision, and F-measure of 80.20%, 60.41%, 86.8%, 60.05%, and 58.61%, respectively. It can be concluded from the results that scary emotions are well discriminated by the architecture, followed by amusing, relaxing, and boring emotions. Studies using the CASE have also reported that the classifiers could discriminate scary emotions better than any other emotional state [8]. This is due to high phasic activity with more recruited sweat glands in the scary state with the high number of activated eccrine sweat glands triggered by postganglionic sudomotor.

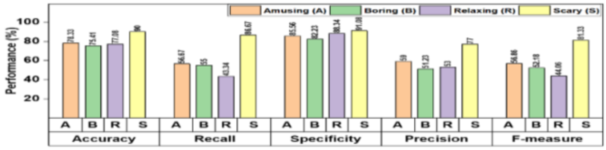


Figure 3. Classification results achieved using cCNN.

4. Limitations and Future Work

The limitations of our study are: Firstly, the use of empirical parameters for TFR and classification architectures. Second, only the CASE dataset is used for training and validation. We must thus evaluate the performance of our architecture across more datasets to assess their generalizability. Third, we have only used EDA and never utilized other physiological modalities like EEG, BVP, and SKT or neurological signals like EMG, which may possess intricate details that can help improve the performance of the classifier [6]. Finally, in actual use, signals from consumer-grade wearables may be lossy for various reasons, including user motions, software issues, and sensor malfunctions.

5. Conclusions

In this study, a process pipeline is proposed for classifying four distinct emotional states using EDA. We analyzed the performance of STFT based TF method and cCNN-based DL architecture in predicting varied human emotions. The results indicate that the proposed STFT-based TFR could better represent the non-stationary EDA signal fluctuations. The cCNN architecture achieved the highest average classification accuracy of 90% in identifying scary emotions with fewer learnable parameters. Thus, the proposed methodology can help diagnose various disorders associated with mental health, such as stress, anxiety, and depression.

References

[1] Ganapathy N, Veeranki YR, Kumar H, Swaminathan R. Emotion recognition using electrodermal activity signals and multiscale deep convolutional neural network. Journal of Medical Systems. 2021 Apr;45(4):1-0.

[2] Institute of Health Metrics and Evaluation. Global Health Data Exchange (GHDx), (https://vizhub.healthdata.org/gbd-results/, accessed 14 May 2022).

[3] Khare SK, Bajaj V. Time±frequency representation and convolutional neural network-based emotion recognition. IEEE Transactions on neural networks and learning systems. 2020 Jul 31;32(7):2901-9.

[4] Sharma K, Castellini C, van den Broek EL, Albu-Schaeffer A, Schwenker F. A dataset of continuous affect annotations and physiological signals for emotion analysis. Scientific data. 2019 Oct 9;6(1):1-3.

[5] Greco A, Valenza G, Lanata A, Scilingo EP, Citi L. cvxEDA: A convex optimization approach to electrodermal activity processing. IEEE Transactions on Biomedical Engineering. 2015 Aug 28;63(4):797- 804.

[6] Karthick PA, Navaneethakrishna M, Punitha N, Fredo AJ, Ramakrishnan S. Analysis of muscle fatigue conditions using time-frequency images and GLCM features. Current Directions in Biomedical Engineering. 2016 Sep 1;2(1):483-7.

[7] Sriram Kumar P, Govarthan PK, Ganapathy N, Jac Fredo AR. Electrodermal activity based analysis of emotion recognition using temporal-morphological features and machine learning algorithms. Journal of Mechanics in Medicine and Biology. 2023 Apr 7.

[8] Hinduja S, Kaur G, Canavan S. Investigation into Recognizing Context Over Time using Physiological Signals. In2021 9th International Conference on Affective Computing and Intelligent Interaction (ACII) 2021 Sep 1 (pp. 1-8). IEEE.