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

doi:10.3233/SHTI230422

Automated Emotion Recognition System Using Blood Volume Pulse and XGBoost Learning

Lokesh Naidu LEBAKA[a1], Sriram Kumar P[a], Praveen Kumar GOVARTHAN[a], Priya Rani[b], Nagarajan GANAPATHY[c] and Jac Fredo AGASTINOSE RONICKOM[a]
[a] *School of Biomedical Engineering, Indian Institute of Technology (BHU), Varanasi, Uttar Pradesh, India*
[b] *Electrical and Biomedical Engineering, School of Engineering, RMIT University, Melbourne, Australia*
[c] *Department of Biomedical Engineering, Indian Institute of Technology Hyderabad, Kandi, Telangana, India*
ORCiD ID: Lokesh Naidu Lebaka https://orcid.org/0000-0002-2990-5474

Abstract. In this study, a new method for detecting emotions using Blood Volume Pulse (BVP) signals and machine learning was presented. The BVP of 30 subjects from the publicly available CASE dataset was pre-processed, and 39 features were extracted from various emotional states, such as amusing, boring, relaxing, and scary. The features were categorized into time, frequency, and time-frequency domains and used to build an emotion detection model with XGBoost. The model achieved the highest classification accuracy of 71.88% using the top 10 features. The most significant features of the model were computed from time (5 features), time-frequency (4 features), and frequency (1 feature) domains. The skewness calculated from the time-frequency representation of the BVP was ranked highest and played a crucial role in the classification. Our study suggests the potential of using BVP recorded from wearable devices to detect emotions in healthcare applications.

Keywords. Emotion detection, blood volume pulse, feature extraction, XGBoost

1. Introduction

Emotions encompass a range of psychological and physiological reactions to events, objects, and situations. Emotion detection is a critical technique used to identify, interpret, and classify an individual's emotional state. This technique has significant implications in the field of medicine, particularly for individuals with psycho-neural disorders, learning disabilities, and an autism spectrum disorder. Emotions can be detected using verbal, non-verbal, and physiological signals [1]. However, verbal and

[1] Corresponding Author: L.N. Lebaka, Lebaka.lokeshnaidu.email: bme20@itbhu.ac.in

non-verbal techniques suffer from a lack of accuracy due to their susceptibility to individual alterations. In contrast, physiological signals guarantee reliability because they are largely involuntarily activated by the autonomic and somatic nervous systems [2]. One of the most widely used physiological signals for emotion detection is the Blood Volume Pulse (BVP) signal, and as it is recorded through a wearable photoplethysmography device, it causes negligible discomfort, hence avoiding alteration of the patient's mood while recording. The optical sensor in these devices measures changes in light absorption density of the skin and tissue, allowing the detection of variations in BVP amplitude that represent instantaneous sympathetic activation through sympathetic and parasympathetic systems of the autonomic nervous system [3].

Machine learning (ML) algorithms help identify and categorize the features corresponding to different emotional states based on predefined training information. In this study, the time, frequency, and time-frequency domain features extracted from the BVP and XGBoost classifier were used to detect the corresponding emotional states. Emotion detection through physiological signals, especially the BVP, has the potential to revolutionize the field of medicine by aiding in the diagnosis and treatment of neuro-developmental disorders such as autism spectrum disorder.

2. Materials and methods

The process pipeline followed in the study is shown in figure 1.

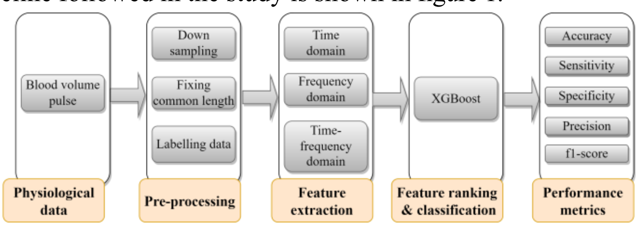


Figure 1. Process pipeline of the proposed study.

The BVP signals considered in our study were obtained from the publicly available Continuously Annotated Signals of Emotion (CASE) dataset [4]. It included physiological signals recorded from 15 male and 15 female subjects (aged 22 to 37). The participants watched eight emotional audio-video clips to elicit the four specific emotions (amusing, boring, relaxing, and scary). Initially, the signal was pre-processed in MATLAB, and the physiological data of the BVP was down sampled to 20 Hz to reduce memory and computation costs. A common and fixed length (from the last part of the recording) is chosen for the analysis of each signal as it contained significant information [5]. Feature extraction was performed on the pre-processed signal. A total of 39 features were extracted, out of which 15 were time-domain features [6], 6 frequency domain features [7], and 18 time-frequency features comprising of Short Time Fourier-Transform (STFT) and Mel-Frequency Cepstral Coefficient (MFCC) features as shown in Table 1.

The features extracted from the BVP were fed to the XGBoost classifier for the feature ranking and emotion classification [8]. The features were normalized (0 to 1) and manually split into five folds for cross-validation, with the data balanced during

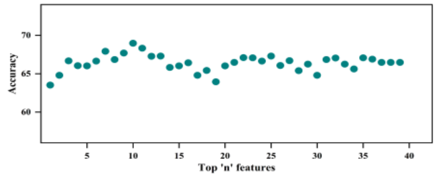
Table 1. List of the features extracted from the BVP.

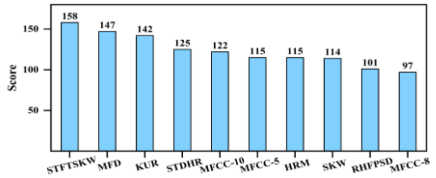
Domain	Features
Time	Mean, Standard deviation, Energy, Arc length, Skewness (SKW), Kurtosis (KUR), Entropy, Mean of First Derivative (MFD), First derivative standard deviation, First derivative mean normalized, Second derivative mean, Second derivative standard deviation, Second derivative mean normalized, Heart Rate Mean (HRM), Standard Deviation of Heart Rate (STDHR)
Frequency	High-frequency power spectral density, Relative High-Frequency Power Spectral Density (RHFPSD), Low-frequency power spectral density, Relative low-frequency power spectral density, Very low-frequency power spectral density, Relative very low-frequency power spectral density
Time-frequency	STFT mean, STFT standard deviation, STFT variance, STFT Skewness (STFTSKW), STFT kurtosis, MFCC features

both training and testing. The 'GridsearchCV' package was used for tuning the hyperparameters ("n estimators", "max depth," and "reg lambda") of the XGBoost model. Features were ranked using XGBoost feature ranking. Further, the performance of the ML model was evaluated for top n features ($1 \leq n \leq 39$) using measures such as accuracy, sensitivity, specificity, precision, and f1-score.

3. Results and discussion

Figure 2(a) shows the performance of the model for different numbers of features ranked using the XGBoost feature ranking method. It can be noted that the accuracy of the classifier was high with the top 10 features. Fewer features led to low accuracy due to missing patterns, while more features caused low accuracy due to the presence of unnecessary patterns. It can be noted that an ML model with an optimal number of features can lead to the best performance. We achieved an average classification accuracy of 68.97%. The corresponding average sensitivity, specificity, precision, and f1-score for these top 10 features were observed to be 37.9%, 79.3%, 37.9%, and 36.8%, respectively. We achieved the highest classification accuracy of 71.87% for emotion detection. Figure 2(b) shows the top 10 features that significantly contributed to the classification model. It can be seen that "STFTSKW" contributed significantly to the classifier.





(a) (b)

Figure 2. (a) Performance of XGBoost for a different set of features (b) Top 10 features ranked according to XGBoost feature ranking method.

4. Limitations and future research

Our study has achieved reasonable performance in the classification of emotional states; however, a few pitfalls were observed in the analysis. We have considered only 30 subjects (CASE) in our analysis; including more samples from other datasets may improve the classification performance. More advanced time, frequency, and time-frequency domain features can be included for in-depth analysis, such as continuous wavelet transform, slope sign changes, etc. We have attempted with the XGBoost classifier for the feature ranking and classification, however the performance can be improved by building the classification models with advanced ML algorithms and deep learning approaches. There is also much scope to improve the performance of our system by incorporating other physiological modalities such as electrodermal activity [9], skin temperature, and electromyography.

5. Conclusion

In this study, a novel approach to emotion recognition has been proposed using the BVP signals of the subjects. We used time, frequency, and time-frequency domain features to classify the four emotions: amusing, boring, relaxing, and scary. We ranked the features and built the classification model using the XGBoost classifier. The proposed framework achieved the highest classification accuracy of 71.87% in classifying the considered emotions using the top 10 features. We found that "STFTSKW" has a vital role in emotion classification. Our results show that the BVP can be used as a sensing modality for emotion detection in wearable devices.

References

[1] Saurabh Singh 2015, EDBL - Algorithm for Detection and Analysis of Emotion Using Body Language.

[2] Guanghao Yin, Shouqian Sun, Hui Zhang, Dian Yu, Chao Li, Kejun Zhang, Ning Zou User independent Emotion Recognition with Residual Signal-Image Network, arXiv:1908.03692

[3] Bota PJ, Wang C, Fred AL, Da Silva HP. A review of current challenges, and future possibilities on emotion recognition using machine learning and physiological signals. IEEE Access. 2019 Sep 26;7:140990- 1020.

[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] Veeranki YR, Ganapathy N, Swaminathan R. Analysis of Fluctuation Patterns in Emotional States Using Electrodermal Activity Signals and Improved Symbolic Aggregate Approximation. Fluctuation and Noise Letters. 2022 Apr 29;21(02):2250013.

[6] Gouizi K, Reguig FB, Maaoui C. Analysis of physiological signals for emotion recognition. International Workshop on Systems, Signal Processing and their Applications, WOSSPA 2011 July 9 (pp. 147-150). IEEE.

[7] Yin Z, Zhao M, Wang Y, Yang J, Zhang J. Recognition of emotions using multimodal physiological signals and an ensemble deep learning model. Computer methods and programs in biomedicine. 2017 Mar 1;140:93-110.

[8] Abirami, S., Thomas, J., Yuvaraj, R., Agastinose Ronickom, J.F. (2022). A Comparative Study on EEG Features for Neonatal Seizure Detection. In: Murugappan, M., Rajamanickam, Y. (eds) Biomedical Signals Based Computer-Aided Diagnosis for Neurological Disorders. Springer, Cham.

[9] Khan AM, Lawo M. Recognizing emotion from blood volume pulse and skin conductance sensor using machine learning algorithms. InXIV Mediterranean Conference on Medical and Biological Engineering and Computing 2016: MEDICON 2016, March 31st-April 2nd, 2016, Paphos, Cyprus 2016 (pp. 1297-1303). Springer International Publishing.