EEG-Based Emotion Recognition via Fast and Robust Feature Smoothing

Cheng Tang[1], Di Wang[2], Ah-Hwee Tan[1,2], and Chunyan Miao[1,2]

[1] School of Computer Science and Engineering
[2] Joint NTU-UBC Research Center of Excellence in Active Living for the Elderly
Nanyang Technological University, Singapore 639798
tang0314@e.ntu.edu.sg, {wangdi,asahtan,ascymiao}@ntu.edu.sg

Abstract. Electroencephalograph (EEG) signals reveal much of our brain states and have been widely used in emotion recognition. However, the recognition accuracy is hardly ideal mainly due to the following reasons: (i) the features extracted from EEG signals may not solely reflect one's emotional patterns and their quality is easily affected by noise; and (ii) increasing feature dimension may enhance the recognition accuracy, but it often requires extra computation time. In this paper, we propose a feature smoothing method to alleviate the aforementioned problems. Specifically, we extract six statistical features from raw EEG signals and apply a simple yet cost-effective feature smoothing method to improve the recognition accuracy. The experimental results on the well-known DEAP dataset demonstrate the effectiveness of our approach. Comparing to other studies on the same dataset, ours achieves the shortest feature processing time and the highest classification accuracy on emotion recognition in the valence-arousal quadrant space.

Keywords: Emotion recognition, EEG, DEAP, Feature smoothing

1 Introduction

Emotion is the subjective experience that reflects our mental states and can significantly affect our cognitive function and action tendencies [10]. With the advances in artificial intelligence (AI) and brain-computer interface (BCI) technologies, the ability for computer applications to recognize human emotions can provide us more intelligent services, such as style-adjusting e-learning system [1], driver's fatigue detection [6], e-healthcare assistance [4], etc.

In efforts to recognize human emotions using machines, researchers mainly rely on the following three types of data: (i) behavioral patterns such as facial expressions, (ii) physiological signals from peripheral nervous system such as electrooculography (EOG), and (iii) physiological signals from central nervous system such as electroencephalograph (EEG). Compared to the other two types of signals, EEG is more informative for high-level brain activities [7]. Moreover, studies showed that EEG exhibits promising characteristics in revealing the subject's emotional states [17]. Thanks to the emerging non-invasive brain-computer interfacing devices, EEG has become one of the most prevalent signals being used to recognize human emotions.

With the aid of machine learning algorithms, many prior studies used different features extracted from raw EEG signals to decode the underlying emotion. However, the recognition accuracy is hardly ideal because of the following reasons: (i) EEG is a mixture of fluctuations induced by many neuronal activities in the brain and is susceptible to interference [13]. The features extracted from EEG may vary drastically within short periods but human emotions are relatively stable, which means the features may not directly reflect the emotional patterns; and (ii) increasing the feature dimension may improve the recognition accuracy, but this approach usually introduces more computational cost in feature extraction, classifier training, and the classification task.

To address the aforementioned feature instability problem without increasing the feature dimension, in this paper, we propose a fast and robust feature smoothing method, which can be applied on the extracted EEG features to improve the emotion recognition accuracy.

Without feature smoothing, emotion-irrelevant patterns make the extracted features less distinctive, thus reduce the classification accuracy. This problem can be alleviated by applying moving average smoothing on the extracted EEG features. This simple feature smoothing method will not increase the feature dimension nor add in much time to the total feature processing process. We evaluate our proposed approach on the widely studied DEAP dataset [8]. Specifically, we extract six statistical features from the EEG signals and apply moving average smoothing on all the extracted features. Using support vector machine (SVM) as the classifier, we obtain 82.3% accuracy in recognizing four classes of emotions, which is higher than four prior studies using the same dataset. Moreover, the processing time of our feature set is the shortest.

The rest of this paper is organized as follows. In Section 2, we review related work on EEG-based emotion recognition. In Section 3, we present the motivation of feature smoothing and the details of our methodology. In Section 4, we show our experimental results on the DEAP dataset with comparisons and discussions. Finally, we conclude and propose future work in Section 5.

2 Related Work

Various prior studies have been conducted to explore how to extract better feature sets for EEG-based emotion recognition. Some studies investigated the characteristics of EEG signals in the frequency domain. Heraz and Frasson [5] used the amplitude of four frequency bands to obtain an averaged accuracy of 74%, 74% and 75% on 17 subjects in the valence, arousal and dominance dimensions, respectively. Bos [2] explored arithmetic combinations of the power on frequency bands to obtain the highest accuracy of 92% in both arousal and valence dimensions on five subjects. Some studies investigated other feature sets such as discrete wavelet coefficients (84.67% for happy and sad on five subjects) [23], fractal dimension (around 90% for arousal on twelve subjects) [16], and higher order crossing (50.13% for four emotions on DEAP dataset) [9]. However, majority of the afore-reviewed results are based on binary emotion classifications, which may not be enough to capture the emotional variations in our daily life. Moreover, as a general observation from the literature, the recognition accuracy decreases, sometimes significantly, if the models need to recognize more classes of emotion. Vyzas et al. [12] managed to recognize six emotions with a remarkable accuracy of 81% on a single subject, but they incorporated other physiological signals such as blood pressure and heart rate besides EEG.

The difficulty in accurately recognizing different emotions merely from EEG signals lies in the non-stability of EEG features. It has been found that feature smoothing can reduce such non-stability and improve recognition accuracy. Shi and Lu first proposed a linear dynamical system (LDS) approach to estimate the latent states of vigilance [15] and later used this model for feature smoothing in emotion recognition (91.77% for positive/negative) [21]. Although LDS is effective in enhancing recognition accuracy, the expectation-maximization (EM) algorithm incorporated in the smoothing process makes the overall approach computationally expensive. Pham et al. [11] used the Savitzky-Golay method, which is based on local least-squares polynomial approximation, to smooth EEG features. Although their proposed method improved the recognition accuracy in the valence dimension, the improved accuracy of 77.38% is not high among similar binary classification problems. In this paper, we apply moving average feature smoothing on six statistical features extracted from the raw EEG signals. As such, the smoothing method does not introduce much computational cost.

3 Moving Average Smoothing on Statistical Feature Set

We adopt the common process to recognize human emotions using EEG signals, i.e., extract features from the raw data, then train the classifier to perform emotion recognition. In addition, we apply feature smoothing on the extracted features before training the classifier. We introduce each of these key steps of our proposed approach with details in the following subsections.

3.1 Feature Extraction

In our proposed approach, we only extract six statistical features, which have been widely adopted in prior studies. Vyzas et al. [18] showed that these six features are strongly correlated to emotions. Moreover, it is computationally inexpensive to extract these simple statistical features.

Let X_n denote an EEG signal value at the nth time stamp, where $n = 1, 2, \ldots, N$ and N denotes the total number of data samples. Moreover, let $\bar{X}_n$ denote the corresponding normalized signal with zero mean and unit variance. Then we extract the following six statistical features:

1. μ, mean of the raw signal over time N:

$$\mu = \frac{1}{N} \sum_{n=1}^{N} X_n. \tag{1}$$

2. σ, standard deviation of the raw signal:

$$\sigma = \sqrt{\frac{1}{N-1}\sum_{n=1}^{N}(X_n - \mu)^2}. \tag{2}$$

3. δ, mean of the absolute values of the first differences of the raw signal:

$$\delta = \frac{1}{N-1}\sum_{n=1}^{N-1}\left|X_{n+1} - X_n\right|. \tag{3}$$

4. $\bar{\delta}$, mean of the absolute values of the first differences of the normalized signal:

$$\bar{\delta} = \frac{1}{N-1}\sum_{n=1}^{N-1}\left|\bar{X}_{n+1} - \bar{X}_n\right| = \frac{\delta}{\sigma}. \tag{4}$$

5. γ, mean of the absolute values of the second differences of the raw signal:

$$\gamma = \frac{1}{N-2}\sum_{n=1}^{N-2}\left|X_{n+2} - X_n\right|. \tag{5}$$

6. $\bar{\gamma}$, mean of the absolute values of the second differences of the normalized signal:

$$\bar{\gamma} = \frac{1}{N-2}\sum_{n=1}^{N-2}\left|\bar{X}_{n+2} - \bar{X}_n\right| = \frac{\gamma}{\sigma}. \tag{6}$$

3.2 Moving Average Smoothing on Extracted Features

Within short time periods, the emotional states of human are relatively stable, but the features obtained from EEG signals may have strong variation in time due to the impact of emotion-irrelevant activities and random fluctuations [13]. To make the features more robust for emotion recognition, we propose to use the moving average method to smooth the features in time sequence. Specifically, we first divide EEG data into non-overlapping windows and extract features from each window. Let f_i denote a single feature f extracted from the ith time window, where $i = 1, 2, \ldots, I$ and I denotes the total number of the non-overlapping windows. The smoothed feature $\bar{f}_i$ is then computed as follows:

$$\bar{f}_i = \frac{1}{T}\sum_{\lfloor i-2/T \rfloor}^{\lfloor i+2/T \rfloor} f_i, \tag{7}$$

where T is the size of the moving average smoother.

3.3 Classification Algorithm

In this paper, we use SVM as the classifier due to its well-known generalization property. In particular, we use the one-vs-all scheme for multiclass classification of the LIBSVM package [3].

For binary classification, given training samples $\{x_i, y_i\}$, where $i = 1, 2, \ldots, l$, $x_i \in \mathbb{R}^d$ and $y_i \in \{-1, 1\}$, SVM solves the following optimization problem:

$$\begin{aligned} \underset{w,b,\xi}{\text{minimize}} \quad & \frac{1}{2}||w||^2 + C\sum_{i=1}^{l}\xi_i, \\ \text{subject to} \quad & y_i(w^T\phi(x_i) + b) \geq 1 - \xi_i, \;\; \xi_i \geq 0, \end{aligned} \tag{8}$$

where C denotes the cost parameter indicating the penalty of error and ξ_i denotes the tolerance of error. Kernel function $\phi(x_i)$ maps the feature vector x_i into another feature space. In this paper, we use radial basis function (RBF) kernel, which is represented as:

$$K(x_i, x_j) \equiv \phi(x_i)^T\phi(x_j) = \exp(-\gamma||x_i - x_j||^2), \gamma > 0, \tag{9}$$

where γ defines the steepness of the decision boundary.

4 Emotion Recognition on DEAP Dataset

To assess the performance of our feature extraction and feature smoothing strategy, we use the well-known DEAP dataset for evaluations. DEAP dataset was collected by Koelstra et al. [8] for human emotion analysis. EEG signals of 32 subjects were elicited using multimodal stimuli and recorded on 32 channels using the Biosemi ActiveTwo system[1]. In the preprocessed dataset provided[2], each subject has 40 minutes' recordings of EEG signals. Moreover, ratings of valence, arousal and dominance were labeled by the subjects after each trial. The EEG data were down-sampled to 128 Hz, filtered by a bandpass filter of 4-45 Hz, and normalized with respect to the common reference in each channel. In this paper, we take the subjects' labels in the valence and arousal dimensions as the ground truth of the EEG data. Actually, we are following the circumplex model of affect proposed by Russel [14]. In his widely adopted model (e.g., applied in [8], [20] and [19]), emotions are represented in a two-dimensional space, where the two axes represent valence and arousal, respectively.

4.1 Experimental Setup

In our experiments, we segment all EEG data given in the DEAP dataset into non-overlapping windows of one second, where each window consists of 128 data samples. Therefore, the total number of observations/windows of a subject is

[1] http://www.biosemi.com

[2] http://www.eecs.qmul.ac.uk/mmv/datasets/deap/

2400 (40 videos times 60 seconds). Moreover, for every observation, we extract the six statistical features from each EEG channel. Because data in DEAP were collected using a 32-channel device, the size of our feature set is 192.

In DEAP , the ratings of valence and arousal were given as decimals in the $[1, 9]$ interval. Therefore, we choose 5 as the threshold for class labeling in the valence-arousal space. In other words, we use the ratings provided by the subjects in DEAP as the ground truth to define four classes of emotion for assessments. These four classes are $V_L A_L$, $V_L A_H$, $V_H A_L$ and $V_H A_H$, where V denotes valence, A denotes arousal, L denotes low value (< 5), and H denotes high value (≥ 5).

To split the training and testing samples, we did not choose the k-fold cross-validation scheme because feature smoothing should only be applied on continuous time sequence that segmenting a continuous feature sequence into k parts will break down the continuity. Instead, we choose 80%/20% splitting strategy to preserve the continuity in features. Specifically, for samples in each minute, we use the first 80% for training and the rest 20% for testing. Feature smoothing is then applied separately on the two sets. This splitting strategy is depicted in Fig. 1(a). Furthermore, we normalize training samples (referring to the extracted features rather than the raw signals) to zero-mean and unit variance to train the classifier, and then use the normalization parameters obtained from training samples to normalize testing samples before performing classification.

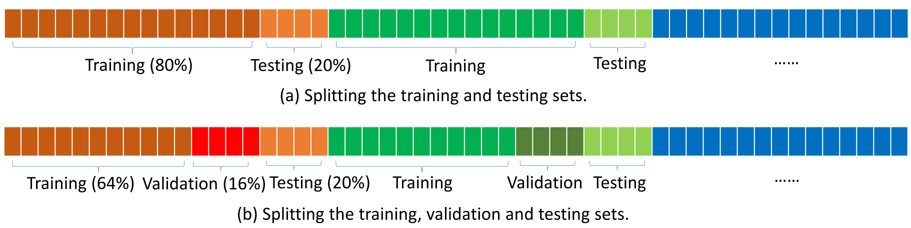

Fig. 1. 80%/20% splitting strategy for obtaining training, validation and testing sets.

In feature smoothing, the window size T (see (7)) greatly affects the performance. To obtain the best value of T, we further split a validation set from the training set using the same splitting ratio as illustrated in Fig. 1(b). Table 1 shows the accuracy obtained on the validation set based on classifiers trained using the training set (64%) with respect to different T values. As shown in Table 1, we obtain the best accuracy when $T = 11$. As such, we use $T = 11$ during subsequent feature smoothing on both the training (80%) and testing datasets.

Table 1. Classification accuracy on the validation set with varying T values

T	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Acc (%)	59.83	68.23	73.55	76.15	77.86	78.96	80.36	80.58	81.13	81.06	**81.31**	80.81	80.77	80.67	80.90	79.49

On the other hand, to find the appropriate parameter settings of SVM, we perform grid search on both the cost parameter C (see (8)) and the penalty parameter γ (see (9)) for $\{n \times 10^m\}$, where $n = 1, 2, \ldots, 9$ and $m = -3, -2, \ldots, 3$. When $T = 11$, we find the best performing values as $C = 10$ and $\gamma = 0.005$. This combination of parameter values are used in all the experiments presented in this paper.

4.2 Results and Discussions

Fig. 2 shows the four-class emotion recognition results of each individual subject before and after applying our moving average feature smoothing method. The average accuracy (across all subjects) improves significantly from $61.73\% \pm 7.07\%$ to $82.30\% \pm 8.44\%$ after applying feature smoothing. This finding strongly support the high effectiveness of our feature smoothing method.

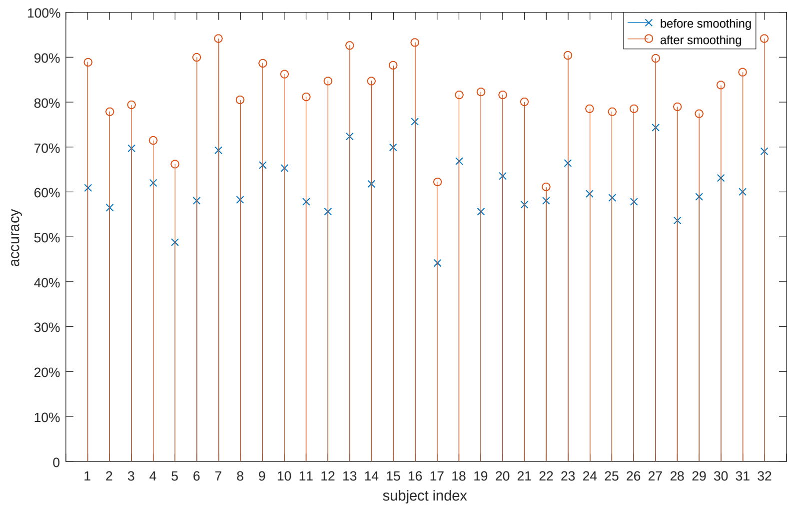

Fig. 2. Emotion recognition accuracy before and after feature smoothing.

Besides 80%/20%, we also conduct experiments on other splitting ratios and then present the results in Table 2. Similar to Fig. 1(b), for each splitting ratio R, we further split a validation set from the training set based on R to obtain the corresponding optimal T value. As clearly shown in Table 2, our feature smoothing method always improves the recognition accuracy. It is more encouraging to see that even when $R = 20\%/80\%$, our approach can obtain an accuracy of 62.43% on four-class emotion recognition. This satisfactory performance obtained on low splitting ratio well demonstrates the generalization capability or robustness of our approach. To be more elaborate, for a longitudinal study with stable emotion transitions, say each session lasts one minute (same as DEAP), our approach only requires the EEG signals collected in the first twelve seconds to be labeled for training, hereafter, it can already achieve 62.43% accuracy on four-class emotion recognition.

Table 2. Performance of feature smoothing using different train/test splitting ratios

Train/Test (%)	Acc before smoothing (%)	T value	Acc after smoothing (%)	Acc improvement (%)
20/80	51.37	5	62.43	11.06
40/60	55.67	7	72.41	16.74
60/40	58.26	11	78.36	20.10
80/20	61.73	11	82.30	20.57

Table 3. Comparison with prior studies on DEAP

Study	Feature Set	Feature Smoothing	Time (s)	Classifier	Subjects	Performance
[8]	PSD (32 channels)+ difference between 14 pairs of symmetric channels	None	166.2	Gaussian Naive Bayes	All 32 subjects	Valence: 57.6%, arousal: 62%
[9]	Six statistics + FD + HOC (32 channels)	None	>1000	SVM with polynomial kernel	All 32 subjects	Four emotions: 80%
[22]	PSD (16 channels)	None	82.6	CNN	22 selected subjects	Valence: 76.63%
[11]	PSD (32 channels)	Savitzky-Golay	207.4	SVDD	All 32 subjects	Three levels of valence: 71.75%
Our work	Six statistics (32 channels)	None	**67.8**	SVM with RBF kernel	All 32 subjects	Valence-arousal quadrant: 61.73%
		Moving average	**77.4**	SVM with polynomial kernel		Valence-arousal quadrant: 67.90%
		Moving average	**77.4**	SVM with RBF kernel		Valence-arousal quadrant: **82.30**%

Note: PSD denotes power spectrum density, FD denotes fractal dimension, HOC denotes higher order crossing, CNN denotes convolutional neural network, and SVDD denotes support vector data description.

To further assess the performance of our proposed method, we compare our results with some prior studies using the same DEAP dataset. In Table 3, [8] used leave-one-video-out scheme for classification, [22] conducted 11-fold cross-validation on 22 selected subjects, [9] and [11] used 5-fold cross-validation. Although our data splitting strategy is different from all the benchmarking studies (similar but conducted four times less than 5-fold CV) and some models adopted different number of classes for emotion recognition, we still compare all the results in the same table. Nonetheless, we use the same computer (2.20 GHz CPU with 8 GB RAM) and the same programming language (MATLAB) to obtain all the feature sets shown in Table 3 and report their processing time (feature extraction and smoothing if applicable). It is encouraging to see that our approach achieves the highest accuracy of 82.3%, which is even higher than those recognizing lesser number of emotional classes. Moreover, our moving average approach does not add in much computational time to the overall feature processing procedure ($(77.4 - 67.8)/77.4 = 12.4\%$). Compared to other benchmarking models, either with or without feature smoothing, our approach has the shortest feature processing time.

5 Conclusion

In this paper, we propose a fast and robust EEG-based emotion recognition approach that applies the simple moving average feature smoothing method on the six extracted statistical features. To assess the effectiveness of our approach, we apply it on the well-known DEAP dataset to perform emotion recognition in the valence-arousal quadrant space. The results are more than encouraging. First of all, the average accuracy (when using 80%/20% splitting ratio) significantly improves from 61.73% to 82.3% after feature smoothing is applied. Secondly, we show the robustness of our approach that it always significantly improves the recognition accuracy for various data splitting ratios. Last but most importantly, our approach achieves the best performance in terms of both feature processing time and recognition accuracy among all the benchmarking models.

In the future, we will further test the robustness of our approach by conducting experiments on own collected and other datasets. We will also look into the theoretical insights of why the simple moving average method may significantly improve the emotion recognition accuracy.

Acknowledgment

This research is supported by the National Research Foundation, Prime Ministers Office, Singapore under its IDM Futures Funding Initiative. This research is also partially supported by the NTU-PKU Joint Research Institute, a collaboration between Nanyang Technological University and Peking University that is sponsored by a donation from the Ng Teng Fong Charitable Foundation.

References

1. Asteriadis, S., Tzouveli, P., Karpouzis, K., Kollias, S.: Estimation of behavioral user state based on eye gaze and head pose-application in an e-learning environment. Multimedia Tools and Applications 41(3), 469–493 (2009)
2. Bos, D.O.: EEG-based emotion recognition-the influence of visual and auditory stimuli. Capita Selecta (MSc course) (2006)
3. Chang, C.C., Lin, C.J.: LIBSVM: A library for support vector machines. ACM Transactions on Intelligent Systems and Technology 2, 27:1–27:27 (2011), software available at `http://www.csie.ntu.edu.tw/~cjlin/libsvm`
4. Goldberg, D.: The detection and treatment of depression in the physically ill. World Psychiatry 9(1), 16–20 (2010)

5. Heraz, A., Frasson, C.: Predicting the three major dimensions of the learners emotions from brainwaves. International Journal of Computer Science 2(3), 187–193 (2007)
6. Jap, B.T., Lal, S., Fischer, P., Bekiaris, E.: Using EEG spectral components to assess algorithms for detecting fatigue. Expert Systems with Applications 36(2), 2352–2359 (2009)
7. Kandel, E.R., Schwartz, J.H., Jessell, T.M., Siegelbaum, S.A., Hudspeth, A.J.: Principles of Neural Science. Mc Graw Hill (2012)
8. Koelstra, S., Muhl, C., Soleymani, M., Lee, J.S., Yazdani, A., Ebrahimi, T., Pun, T., Nijholt, A., Patras, I.: DEAP: A database for emotion analysis using physiological signals. IEEE Transactions on Affective Computing 3(1), 18–31 (2012)
9. Liu, Y., Sourina, O.: EEG databases for emotion recognition. In: International Conference on Cyberworlds. pp. 302–309. IEEE (2013)
10. Mauss, I.B., Robinson, M.D.: Measures of emotion: A review. Cognition and Emotion 23(2), 209–237 (2009)
11. Pham, T.D., Tran, D., Ma, W., Tran, N.T.: Enhancing performance of EEG-based emotion recognition systems using feature smoothing. In: International Conference on Neural Information Processing. pp. 95–102. Springer (2015)
12. Picard, R.W., Vyzas, E., Healey, J.: Toward machine emotional intelligence: Analysis of affective physiological state. IEEE Transactions on Pattern Analysis and Machine Intelligence 23(10), 1175–1191 (2001)
13. Pijn, J.P., Van Neerven, J., Noest, A., da Silva, F.H.L.: Chaos or noise in EEG signals' dependence on state and brain site. Electroencephalography and Clinical Neurophysiology 79(5), 371–381 (1991)
14. Russell, J.A.: A circumplex model of affect. Journal of Personality and Social Psychology 39(6), 1161–1178 (1980)
15. Shi, L.C., Lu, B.L.: Off-line and on-line vigilance estimation based on linear dynamical system and manifold learning. In: International Conference on Engineering in Medicine and Biology. pp. 6587–6590. IEEE (2010)
16. Sourina, O., Liu, Y.: A fractal-based algorithm of emotion recognition from EEG using arousal-valence model. In: BIOSIGNALS. pp. 209–214 (2011)
17. Takahashi, K.: Remarks on emotion recognition from multi-modal bio-potential signals. In: International Conference on Industrial Technology. vol. 3, pp. 1138–1143. IEEE (2004)
18. Vyzas, E., Picard, R.W.: Affective pattern classification. In: Emotional and Intelligent: The Tangled Knot of Cognition. pp. 176–182 (1998)
19. Wang, D., Tan, A.H., Miao, C.: Modelling autobiographical memory in human-like autonomous agents. In: International Conference on Autonomous Agents and Multiagent Systems. pp. 845–853. ACM (2016)
20. Wang, D., Tan, A.H.: Mobile humanoid agent with mood awareness for elderly care. In: International Joint Conference on Neural Networks. pp. 1549–1556. IEEE (2014)
21. Wang, X.W., Nie, D., Lu, B.L.: Emotional state classification from EEG data using machine learning approach. Neurocomputing 129, 94–106 (2014)
22. Yanagimoto, M., Sugimoto, C.: Recognition of persisting emotional valence from EEG using convolutional neural networks. In: International Workshop on Computational Intelligence and Applications. pp. 27–32. IEEE (2016)
23. Yohanes, R.E., Ser, W., Huang, G.B.: Discrete wavelet transform coefficients for emotion recognition from EEG signals. In: International Conference on Engineering in Medicine and Biology. pp. 2251–2254. IEEE (2012)