2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

EEG-based Emotion Recognition Using Similarity Measure of Brain Rhythm Sequencing

Jia Wen Li[1,2,3], Shovan Barma[4], Sio Hang Pun[1,2], Fei Chen[5], Cheng Li[1,2,3,5], Ming Tao Li[1,2,3,5], Pan Ke Wang[1,2,3], Mang I Vai[1,2,3], and Peng Un Mak[3]

***Abstract*—The similarity is a fundamental measure from the homology theory in bioinformatics, and the biological sequence can be classified based on it. However, such an approach has not been utilized for electroencephalography (EEG)-based emotion recognition. To this end, the sequence generated by choosing the dominant brain rhythm owning maximum instantaneous power at each 0.2 s timestamp of the EEG signal has been proposed. Then, to recognize emotional arousal and valence, the similarity measures between pairwise sequences have been performed by dynamic time warping (DTW). After evaluations, the sequence that provides the highest accuracy has been obtained. Thus, the representative channel has been found. Besides, the appropriate time segment for emotion recognition has been estimated. Those findings helpfully exclude redundant data for assessing emotion. Results from the DEAP dataset displayed that the classification accuracies between 72%–75% can be realized by applying the single-channel data with a 5 s length, which is impressive when considering fewer data sources as the primary concern. Hence, the proposed idea would open a new way that uses the similarity measures of sequences for EEG-based emotion recognition.**

***Index Terms*—Electroencephalography (EEG), brain rhythm sequencing (BRS), similarity measure, sequence classification, emotion recognition.**

I. Introduction

Emotion is one of the fundamental psychological factors that can influence many aspects of daily life, including communication skills, social interaction, and work efficiency. Therefore, its automated recognition is meaningful. In recent years, the electroencephalography (EEG) signal exhibits great potential in this field as it records the neural oscillations that appeared in the brain, which shows a closer relationship with the emotional reaction.

Toward EEG-based emotion recognition, the trustworthy feature is a primary concern. In previous works [1]–[3], it is found that the brain rhythms are the related features that have been widely employed. Specifically, they are originated from classifying the EEG into five frequency sub-bands: 0–4 Hz (δ), 4–8 Hz (θ), 8–13 Hz (α), 13–30 Hz (β), and 30–50 Hz (γ) [4]. Moreover, the variations of such brain rhythms have been viewed as the characteristics to recognize emotional arousal and valence. For instance, Koelstra *et al.* [1] explored that the powers of θ, α, and γ show negative correlations with arousal; Kim *et al.* [2] reported that the α power usually changes with the valence states of fear (negative) and happiness (positive). In addition, the frontal asymmetry of α power shows a steady correlate of valence; Onton and Makeig [3] concluded that there is a positive correlation between the powers of high frequency sub-bands (such as β and γ) and valence.

The aforementioned works were based on the properties of rhythmic powers. However, the characteristics concerning the time-related occurrences of particular brain rhythms have not been investigated. Considering the similarity is a fundamental measure from the homology theory in bioinformatics [5], and biological sequence can be classified based on it, nonetheless, such a method has not been utilized for EEG-based emotion recognition, so an approach named brain rhythm sequencing (BRS) that interprets the EEG as the time-related sequential format consists of dominant rhythms has been proposed in this work. By applying it, the time-frequency characteristics of EEG can be presented simultaneously, which are available to perform the sequence classification through the similarity measure. Besides, after evaluating these sequences generated from the EEG recordings on different channels and times, the representative channel that yields the highest accuracy can be determined accordingly. Meanwhile, the time segment that is appropriate for recognition can also be estimated during the assessments. Those properties obtained can helpfully exclude redundant data and save the computation cost for EEG-based emotion recognition. Hence, the proposed idea would open a novel way that adopts the similarity measure of brain rhythm sequence classification to design the portable emotion-aware application when considering fewer data sources as the main concern.

The remainder is arranged as follows: Section II describes the experimental data from the DEAP dataset. Then, Section III elaborates on the proposed methodology. Section IV shows the results and discussion. Finally, the conclusion of this work has been drawn in Section V.

This work was funded by The Science and Technology Development Fund, Macau SAR (File no. 0144/2019/A3, 0022/2020/AFJ, SKL-AMSV (FDCT-funded), SKL-AMSV-ADDITIONAL FUND, SKL-AMSV(UM)-2020-2022), and funded by the University of Macau (File no. MYRG2018-00146-AMSV, MYRG2019-00056-AMSV), and funded by the National Key R&D Program of China (No. 2020YFB1313502).

[1]J. W. Li, S. H. Pun, C. Li, M. T. Li, P. K. Wang, and M. I Vai are with the Institute of Microelectronics, University of Macau, Macau 999078, China. (e-mail: gzcat29@126.com; lodgepun@um.edu.mo; yb77453@um.edu.mo; yb97473@um.edu.mo; ioewpk@gmail.com; fstmiv@um.edu.mo).

[2]J. W. Li, S. H. Pun, C. Li, M. T. Li, P. K. Wang, and M. I Vai are also with the State Key Laboratory of Analog and Mixed-Signal VLSI, University of Macau, Macau 999078, China.

[3]J. W. Li, C. Li, M. T. Li, P. K. Wang, M. I Vai, and P. U. Mak* are with the Department of Electrical and Computer Engineering, University of Macau, Macau 999078, China (*e-mail: fstpum@um.edu.mo).

[4]S. Barma is with the Department of Electronics and Communication Engineering, Indian Institute of Information Technology Guwahati (IIITG), Guwahati 781015, India (e-mail: shovan@iiitg.ac.in).

[5]F. Chen*, C. Li, and M. T. Li are with the Department of Electrical and Electronic Engineering, Southern University of Science and Technology, Shenzhen 518055, China (*e-mail: fchen@sustech.edu.cn).

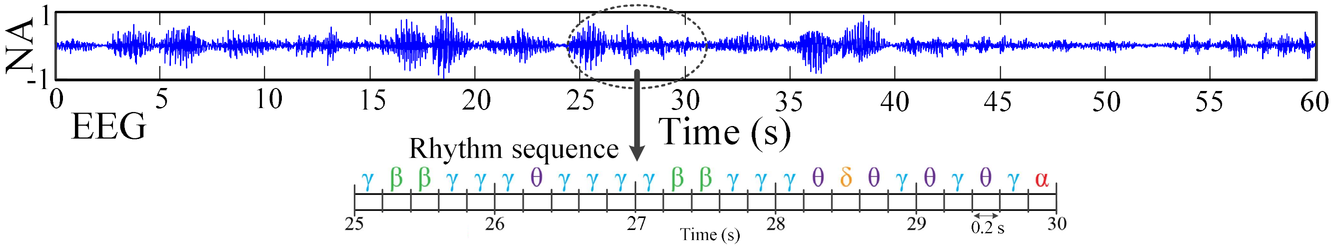

Figure 1. Proposed BRS by choosing the dominant brain rhythm owning maximum power at each 0.2 s timestamp of an EEG signal. (NA: normalized amplitude)

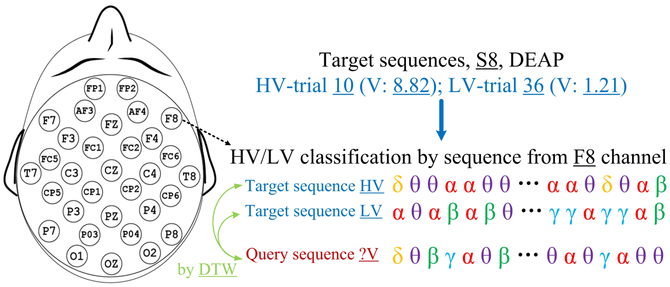

Figure 2. Valence classification (HV/LV) using sequences from the F8 channel of subject S8, DEAP. Here, DTW is used for measuring the similarity, and the query sequence is assigned as the same emotional state as the target sequence that presents a higher similarity level.

II. Experimental Dataset

In this work, the EEG signals for the experiment are from the DEAP dataset [1]. It included 32 subjects (17 males and 15 females, age: 23–37 years). Each subject watched 40 different one-minute long music videos to elicit various emotional states. After watching each video, the subjects rated it (1–9) based on arousal (A) and valence (V). Thus, according to the individual ratings, the emotions have been labeled into several sub-groups for recognition. In arousal: high arousal (HA) with $A \geq 5$ and low arousal (LA) with $A < 5$; in valence: high valence (HV) with $V \geq 5$ and low valence (LV) with $V < 5$. Besides, a 10–20 EEG system involving 32 channels was applied for recordings. So, for each subject, the data size was 60 s × 32 EEG channels × 40 trials. Furthermore, the sampling rate was 128 Hz, and the analog passband filtering with 0.01–100 Hz has been used for data pre-processing, along with removing the artifacts.

III. Methodology

The BRS aims to interpret the EEG in a chronological order based on dominant brain rhythms, which can be accomplished by the time-frequency analysis (TFA). Moreover, the Wigner-Ville distribution (WVD) is a typical TFA category that enables the evaluation of signal power in the specific frequency domain and then localizes it into a corresponding time, which provides vital properties for realizing BRS. However, the WVD usually causes cross-terms that restrict obtaining the frequency information at a shorter instant precisely. Therefore, a suitably smoothing of the WVD along the time and the frequency directions is normally adopted. Such a variant is named the smoothed pseudo WVD (SPWVD) (1):

$$SPW_x(t,\omega) = \int_{-\infty}^{+\infty} h(\tau) \int_{-\infty}^{+\infty} g(s-t)x(s+\frac{\tau}{2})x^*(s-\frac{\tau}{2})e^{-j\omega\tau}dsd\tau \tag{1}$$

The independent control of $h(t)$ and $g(t)$ help to migrate the cross-terms. In addition, to calculate the precise time indices of the higher power regions in the generated time-frequency plane, the reassignment is considered [6]. Its operation is to relocate each value of the SPWVD at any point (t, ω) to another point $(\hat{t}, \hat{\omega})$, which is the center of gravity of the signal power distribution around (t, ω). So, the reassigned value of SPWVD at any point $(\hat{t}, \hat{\omega})$ is the sum of all values reassigned to that point. This approach is reassigned SPWVD (RSPWVD) (2):

$$SPW_x^{(r)}(t',\omega';g,h) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} SPW_x(t,\omega;g,h) \delta(t'-\hat{t}(x;t,\omega))\delta(\omega'-\hat{\omega}(x;t,\omega))dtd\omega \tag{2}$$

where

$$\hat{t}(x;t,\omega) = t - \frac{SPW_x(t,\omega;\tau_g,h)}{2\pi SPW_x(t,\omega;g,h)} \tag{3}$$

$$\hat{\omega}(x;t,\omega) = \omega + j\frac{SPW_x(t,\omega;g,D_h)}{2\pi SPW_x(t,\omega;g,h)} \tag{4}$$

with $\tau_g = tg(t)$ and $D_h(t) = dh(t)/dt$.

Now, aiming to generate the rhythm sequence data from the time-frequency plane, the whole time is divided into several timestamps with each of 0.2 s length, which is decided by the relationship between the average reaction time of neurons and EEG [7]. On the other side, the frequency direction is divided into five parts according to the ranges of five brain rhythms. Undoubtedly, the sequence needs only one dominant rhythm denoted at each timestamp. Hence, the instantaneous rhythmic power is investigated, as the maximum one can reflect the vital sub-band that provides a greater contribution in terms of power discharge, which is pivotal information of the EEG. As a result, the dominant rhythm that shows maximum power contribution at each 0.2 s timestamp is chosen for the BRS. Following this way, Fig. 1 depicts a sample of BRS, in which an EEG signal from the FP1 channel of subject S1 in the DEAP dataset locates at the top, and its rhythm sequence at 25–30 s lies at the bottom.

Next, the generated sequences can be evaluated for emotion recognition based on the similarity measure. Logically, the similar structures between the two sequences indicate similar functions; whereas, the dissimilarities incur various categories. Thus, the similarity measure can be applied for recognizing the rhythm sequences into different emotional states. Moreover, to perform the similarity measure for sequence classification, the target sequences related to specific emotions should be found in advance, which can be regarded as the standard templates. To this end, the individual ratings from the self-assessments are utilized. Specifically, among the 40 experimental trials, the one with the maximum A or V is assigned as HA or HV. Inversely, the minimum A or V is assumed as LA or LV. Besides, if more than one trial is equal to the maximum or the minimum, the trial presented earlier is considered. For example, as for the valence classification of subject S8, the maximum and minimum were 8.82 at trial 10 and 1.21 at trial 36. So, the generated sequences from these two trials have been used as the target sequences of HV and LV, as illustrated in Fig. 2. Then, the other trials have been assigned as query sequences accordingly. Furthermore, dynamic time warping (DTW), a distance-based approach that calculates an optimal alignment is exploited to measure the similarity level between pairwise sequences. Consequently, the query sequence is assigned as the same emotional state as the target sequence that exhibits a higher similarity result. Besides, the sequences are generated based on the channels and times, hence, the evaluations in this work have been performed on 32 channels and 5 types of time lengths (5 s, 10 s, 20 s, 30 s, and 60 s). Finally, the vital properties for emotion recognition, such as the representative channel and appropriate time segment, have been obtained according to the highest classification accuracy.

IV. Results and Discussion

Table I displays the accuracies (mean ± standard deviation) based on different time lengths of sequences from 32 subjects, in which the first column denotes the length and the rest are the accuracies of arousal and valence. As seen, the accuracies by using diverse lengths are close. A short length is helpful to save computation cost, so the 5 s has been chosen to segment the sequences for emotion recognition. In this way, the arousal classification results of subject S13 have been illustrated in Fig. 3, in which the three maps depict the accuracies of 32 channels at the three chosen segments from the periods of start (0–5 s), middle (25–30 s), and end (55–60 s) parts respectively. Here, the deeper the red, the higher the classification accuracy. In Fig. 3, it can approximately disclose the performance of sequence classification during the whole duration, and meanwhile, the accuracy on the same channel dynamically changes during the process. For example, the PO3 channel at 0–5 s provides higher accuracy; whereas, it has lower results at other time segments. It implies that the sequence from the PO3 channel at the start period is more useful to recognize the arousal of S13. So, the representative channel is PO3 and the appropriate time segment is 0–5 s for arousal recognition of S13. Following this way, the results of all 32 subjects have been summarized in Table II.

In Table II, as for the representative channel, besides the two subjects (S1 and S6) are at the same location for arousal and valence classifications, the others are from different locations. It reveals that diverse locations cope with particular dimensions in emotion recognition for most of the cases. Such a statistical result is consistent with [8]. Furthermore, about the appropriate time segments, S18 and S31 are found at the same period for arousal and valence classifications; whereas, the segments of the others are separated. It discloses that different emotional dimensions are usually elicited by various pieces of the stimuli. Meanwhile, it is observed that for 20 subjects, the appropriate time segment of arousal is earlier than valence. It indicates that during the elicitation, arousal has been analyzed preferentially, then valence has been recognized later. In addition, the results of Table II also exhibit the individual characteristics in emotion recognition, as the channels and times are varied among the subjects. It is reasonable because there is a common sense in emotion science that emotional reaction is more relevant to the experiences, backgrounds, and cultures of the subjects [9]. In this regard, the subject-dependent analysis generally produces a better result than the subject-independent method. As a result, the personalized models based on vital characteristics such as the representative channel and appropriate time segment are valuable for achieving EEG-based emotion recognition.

A comparative study with previous related works has been conducted in Table III, in which the first column lists the work, and the rest display the number of channels used, the applied time length of EEG data, main methodology, and the accuracies for arousal and valence classifications correspondingly.

TABLE I

Applied Time Lengths of Sequence and Classification Accuracies

Length (s)	Arousal classification (%)	Valence classification (%)
5	**75.66 ± 6.34**	**72.86 ± 3.30**
10	74.51 ± 6.90	71.63 ± 3.71
20	73.03 ± 6.45	70.07 ± 4.47
30	71.79 ± 7.21	69.65 ± 4.17
60	69.33 ± 7.12	67.19 ± 4.58

TABLE II

Representative Channel and Appropriate Time Segment (5 s Length) for Emotion Recognition of 32 Subjects in The DEAP Dataset

Subject	Arousal classification			Valence classification		
	RC	ATS (s)	ACC (%)	RC	ATS (s)	ACC (%)
S1	FC5	5–10	68.42	FC5	45–50	73.68
S2	F3	0–5	68.42	AF4	5–10	71.05
S3	PZ	15–20	84.21	F7	25–30	65.79
S4	P4	55–60	81.58	CP2	45–50	76.32
S5	PO3	35–40	73.68	P3	30–35	73.68
S6	T7	30–35	71.05	T7	40–45	78.95
S7	P7	40–45	73.68	PZ	0–5	76.32
S8	F3	5–10	71.05	CP1	15–20	71.05
S9	F7	50–55	76.32	CP6	20–25	73.68
S10	T7	20–25	73.68	F3	35–40	68.42
S11	P7	35–40	73.68	F3	50–55	73.68
S12	FP1	10–15	86.84	CZ	15–20	71.05
S13	PO3	0–5	89.47	AF3	10–15	68.42
S14	CP2	40–45	81.58	T8	5–10	68.42
S15	P7	0–5	71.05	C3	25–30	68.42
S16	C4	10–15	73.68	CP2	35–40	81.58
S17	FP2	15–20	73.68	O1	35–40	73.68
S18	FZ	10–15	71.05	T7	10–15	73.68
S19	FZ	0–5	76.32	O2	5–10	71.05
S20	CZ	55–60	84.21	O1	25–30	76.32
S21	OZ	0–5	84.21	FC6	30–35	71.05
S22	PO4	35–40	73.68	FZ	50–55	73.68
S23	AF3	15–20	71.05	T7	45–50	73.68
S24	F3	20–25	84.21	AF4	0–5	73.68
S25	T8	15–20	81.58	PZ	55–60	71.05
S26	FC6	0–5	68.42	FZ	45–50	73.68
S27	F3	55–60	73.68	FC6	10–15	76.32
S28	FZ	40–45	68.42	P3	45–50	71.05
S29	P4	25–30	76.32	AF4	10–15	71.05
S30	FC2	35–40	65.79	P7	0–5	76.32
S31	CZ	35–40	68.42	FZ	35–40	73.68
S32	FC6	15–20	81.58	F3	25–30	71.05

Acronym: RC-representative channel; ATS-appropriate time segment. ACC-accuracy.

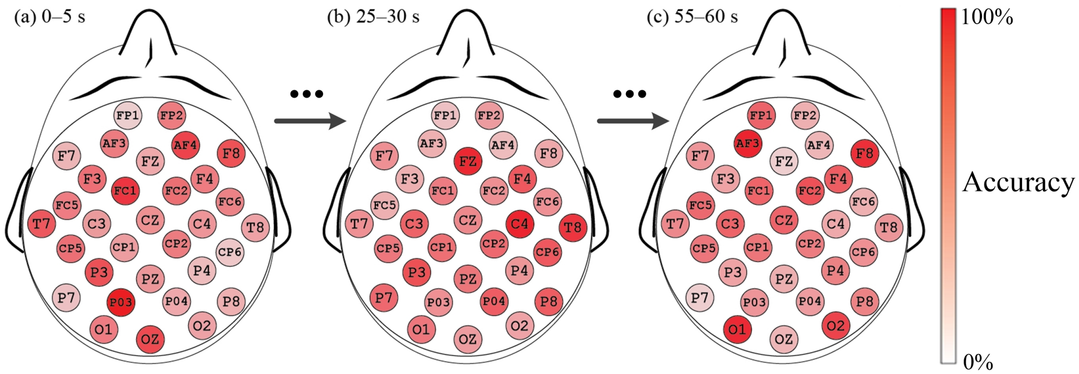

Figure 3. Evaluations of arousal classification (HA/LA) accuracies by using the rhythm sequences from different channels and time segments. The deeper the red, the higher the accuracy. Here, the three chosen segments presented are at the start, middle, and end parts: (a) 0–5 s; (b) 25–30 s; (c) 55–60 s. (subject S13, DEAP)

TABLE III
COMPARATIVE STUDY OF EEG-BASED EMOTION RECOGNITION WORKS

	Number of channels used	**Applied time length of EEG data**	**Main methodology**	**Classification accuracy (%)**	
				Arousal	**Valence**
Chao *et al.* [10]	32	60 s	Capsule network	68.28	66.73
Yoon *et al.* [11]	32	60 s	Bayesian weighted-log-posterior function	70.01	70.09
Atkinson *et al.* [12]	14	60 s	Maximum relevance minimum redundancy method	73.06	73.14
Kumar *et al.* [13]	2	30 s	Bispectral analysis	64.84	61.17
Zhuang *et al.* [14]	8	5 s	Empirical mode decomposition	71.99	69.10
This work	**1**	**5 s**	**BRS with similarity measure**	**75.66**	**72.86**

As seen, the previous related works have not considered the similarity measure of sequence classification. Meanwhile, the single-channel solution for emotion recognition was ignored. The fewer the channels, the fewer the electrodes. Therefore, the representative channel found by the proposed idea is beneficial for designing the portable emotion-aware device. Regarding the applied time length, Zhuang *et al.* [14] utilized 5 s EEG data for recognition. But, which 5 s is proper has not been investigated. So, the appropriate time segment based on 5 s has been assessed in this work. Finally, the proposed method produces accuracies between 72%–75% by employing single-channel and 5 s data only, which is impressive when considering fewer data sources.

V. CONCLUSION

In this work, the BRS with similarity measure has been proposed for EEG-based emotion recognition, and to evaluate its performance, 32 subjects from the DEAP dataset have been studied. Results displayed that the classification accuracies of 72%–75% have been achieved by utilizing fewer data sources compared with the previous works. Besides, the representative channel that discloses vital scalp location, and the appropriate time segment that indicates a short period for recognition, have been investigated. Such properties not only disclose individual characteristics for emotion recognition but also helps to design the emotional-aware device by applying the single-channel data. Hence, the proposed idea would open a novel way that uses the similarity measure of sequence classification for EEG-based emotion recognition.

ACKNOWLEDGMENT

The authors would like to express thanks to the financial supports from the Science and Technology Development Fund of Macau (FDCT), the Research Committee of the University of Macau, and the National Key R&D Program of China.

REFERENCES

[1] S. Koelstra, C. Mühl, M. Soleymani, J. S. Lee, A. Yazdani, T. Ebrahimi, T. Pun, A. Nijholt, and I. Patras, "DEAP: A database for emotion analysis using physiological signals," *IEEE Trans. Affect. Comput*., vol. 3, no. 1, pp. 18–31, 2012.

[2] M. K. Kim, M. Kim, E. Oh, and S. P. Kim, "A review on the computational methods for emotional state estimation from the human EEG," *Comput. Math. Methods Med*., vol. 2013, article 573734, 2013.

[3] J. Onton and S. Makeig, "High-frequency broadband modulation of electroencephalographic spectra," *Front. Hum. Neurosci.*, vol. 3, article 61, 2009.

[4] S. J. Choi and B. G. Kang, "Prototype design and implementation of an automatic control system based on a BCI," *Wirel. Pers. Commun.*, vol. 79, no. 4, pp. 2551–2563, 2014.

[5] J. Zhou, P. Zhong, and T. Zhang, "A novel method for alignment-free DNA sequence similarity analysis based on the characterization of complex networks," *Evol. Bioinforma.*, vol. 12, pp. 229–235, 2016.

[6] S. Barma, B. W. Chen, K. L. Man, and J. F. Wang, "Quantitative measurement of split of the second heart sound (S2)," *IEEE/ACM Trans. Comput. Biol. Bioinf.*, vol. 12, no. 4, pp. 851–860, 2015.

[7] J. W. Li, S. Barma, P. U. Mak, S. H. Pun, and M. I. Vai, "Brain rhythm sequencing using EEG signals: A case study on seizure detection," *IEEE Access*, vol. 7, pp. 160112–160124, 2019.

[8] Y. Dai, X. Wang, X. Li, and Y. Tan, "Sparse EEG compressive sensing for web-enabled person identification," *Measurement*, vol. 74, pp. 11–20, 2015.

[9] N. Lim, "Cultural differences in emotion: differences in emotional arousal level between the East and the West," *Integr. Med. Res.*, vol. 5, no. 2, pp. 105–109, 2016.

[10] H. Chao, L. Dong, Y. Liu, and B. Lu, "Emotion recognition from multiband EEG signals," *Sensors*, vol. 19, article 2212, 2019.

[11] H. J. Yoon and S. Y. Chung, "EEG-based emotion estimation using Bayesian weighted-log-posterior function and perceptron convergence algorithm," *Comput. Biol. Med.*, vol. 43, no. 12, pp. 2230–2237, 2013.

[12] J. Atkinson and D. Campos, "Improving BCI-based emotion recognition by combining EEG feature selection and kernel classifiers," *Expert Syst. Appl.*, vol. 47, pp. 35–41, 2016.

[13] N. Kumar, K. Khaund, and S. M. Hazarika, "Bispectral analysis of EEG for emotion recognition," *Procedia Comput. Sci*., vol. 84, pp. 31–35, 2016.

[14] N. Zhuang, Y. Zeng, L. Tong, C. Zhang, H. Zhang, and B. Yan, "Emotion recognition from EEG signals using multidimensional information in EMD Domain," *Biomed Res. Int.*, vol. 2017, article 8317357, 2017.